

# **Role of small non-coding RNAs in retrograde signalling and stress acclimation in *Arabidopsis thaliana***



## **Dissertation**

zur Erlangung des Grades eines Doktors der  
Naturwissenschaften

an der Fakultät für Biologie

der Ludwig-Maximilians-Universität München

vorgelegt von

**Kristin Habermann**

München, 07.11.2020

1. Gutachter: Prof. Dr. Wolfgang Frank

2. Gutachter: PD Dr. Tatjana Kleine

Tag der Abgabe: 07.11.2020

Tag der mündlichen Prüfung: 04.02.2021

## **STATUTORY DECLARATION AND STATEMENT**

### **Eidesstattliche Erklärung**

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbstständig und ohne unerlaubte Hilfe angefertigt wurde. Des Weiteren erkläre ich, dass ich nicht anderweitig ohne Erfolg versucht habe, eine Dissertation einzureichen oder mich der Doktorprüfung zu unterziehen. Die folgende Dissertation liegt weder ganz, noch in wesentlichen Teilen einer anderen Prüfungskommission vor.

München, 07.11.2020

---

Kristin Habermann

### **Statutory Declaration**

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources. As well, I declare that, I have not submitted a dissertation without success and not passed the oral exam. The present dissertation (neither the entire dissertation nor parts) has not been presented to another examination board.

Munich, 07.11.2020

---

Kristin Habermann

**TABLE OF CONTENT**

<b>LIST OF ABBREVIATIONS.....</b>	<b>2</b>
<b>LIST OF PUBLICATIONS.....</b>	<b>5</b>
<b>DECLARATION OF CONTRIBUTION AS A CO-AUTHOR .....</b>	<b>6</b>
<b>SUMMARY.....</b>	<b>7</b>
<b>ZUSAMMENFASSUNG .....</b>	<b>8</b>
<b>INTRODUCTION.....</b>	<b>9</b>
Endosymbiotic theory .....	9
Anterograde and retrograde signalling .....	10
Biogenesis and functions of small non-coding RNAs .....	15
Biogenesis and functions of long non-coding RNAs .....	21
Non-coding RNAs in biotic and abiotic stress response in <i>Arabidopsis thaliana</i> .....	23
<b>AIM OF THE THESIS .....</b>	<b>27</b>
<b>RESULTS .....</b>	<b>28</b>
Publication I: Impact of small RNAs in retrograde signalling pathways in <i>Arabidopsis thaliana</i> .....	28
Publication II: Identification of cold stress regulated small RNAs in <i>Arabidopsis thaliana</i> .....	<b>Fehler! Textmarke nicht definiert.</b>
Unpublished work: Functional analysis of miRNA overexpression lines in <i>Arabidopsis thaliana</i> .....	<b>Fehler! Textmarke nicht definiert.</b>
<b>DISCUSSION.....</b>	<b>Fehler! Textmarke nicht definiert.</b>
<b>REFERENCES.....</b>	<b>Fehler! Textmarke nicht definiert.</b>
<b>ACKNOWLEDGMENTS .....</b>	<b>Fehler! Textmarke nicht definiert.</b>
<b>CURRICULUM VITAE .....</b>	<b>Fehler! Textmarke nicht definiert.</b>



**LIST OF ABBREVIATIONS**

ABA	abscisic acid
ABFs	ABRE-binding factors
ABF3	ABRE-binding factor3
ABI4	ABSCISIC ACID INSENSITIVE 4
ABRE	ABA-responsive elements
AGO1	ARGONAUTE1
AOX1a	Alternative oxidase 1a
APS	ATP sulfurylase
AREB	ABRE-binding protein
ARF	AUXIN RESPONSE FACTORS
ARLPK	atypical receptor-like pseudokinases
ASCO	the alternative splicing competitor
AtMYB2	MYB DOMAIN PROTEIN 2
bZIP	basic leucine zipper
CBF	C-repeat/drought-responsive element binding factor
CRT	cold response sensitive transcription factors
CSD2	COPPER/ZINC SUPEROXIDE DISMUTASE 2
CSP41b	CHLOROPLAST STEM-LOOP BINDING PROTEIN OF 41 KDA
COLDAIR	COLD ASSISTED INTRONIC NONCODING RNA
COR	cold-responsive
COOLAIR	COLD INDUCED LONG ANTISENSE INTRAGENIC RNAs
DCL1	DICER-LIKE1
DEGs	differentially expressed genes
DRE	dehydration responsive elements
DREB	dehydration responsive element binding
dsRNA	double-stranded RNA
easiRNAs	epigenetically-activated small interfering RNAs
EMS	ethyl methanesulfonate
FDR	false discovery rate
FLC	FLOWERING LOCUS C
GCN2	GENERAL CONTROL NON-REPRESSIBLE 2

gun	GENOME UNCOUPLED
GUS	$\beta$ -glucuronidase
HEN1	HUA ENHANCER 1
hetsiRNAs	heterochromatic siRNAs
hpRNA	hairpin RNA
HSP	heat shock protein
HST	HASTY
HYL1	HYPONASTIC LEAVES 1
IAA28	INDOLE-3-ACETIC ACID INDUCIBLE 28
ICE1	inducer of CBF expression 1
IPS1	INDUCED BY PHOSPHATE STARVATION1
LCR	LEAF CURLING RESPONSIVENESS
LHCB	light harvesting complex of photosystem II
lncRNAs	long non-coding RNAs
LPE1	LOW PHOTOSYNTHETIC EFFICIENCY 1
Mg-Porto	Magnesium-protoporphyrin IX
miRNAs	microRNAs
nat-siRNA	natural antisense transcript-derived siRNA
NB-LRR	nucleotide-binding leucine-rich repeat
ncRNAs	non-coding RNAs
NF	norflurazon
NGS	next generation sequencing
NRPD1A	NUCLEAR RNA POLYMERASE D 1A
NSR	nuclear speckle RNA-binding protein
P5C	DELTA1-PYROLINE-5-CARBOXYLATE
P5CDH	DELTA1-PYRROLINE-5-CARBOXYLATE DEHYDROGENASE
PAC	PALE CRESS
PAP	3'-phosphoadenosine 5'-phosphate
PAP1	PRODUCTION OF ANTHOCYANIN PIGMENT 1
PEP	plastid encoded RNA polymerase
PhANGs	photosynthesis associated nuclear genes
phasiRNAs	phased small interfering RNAs

PHT1	inorganic phosphate transporters
PHO2	UBIQUITIN-CONJUGATING ENZYME 24
Pi	inorganic phosphate
piRNAs	piwi-interacting RNA
PPR	pentatricopeptide repeat
pre-miRNAs	precursor microRNAs
pri-miRNA	primary microRNAs
Proto	protoporphyrin IX
PSBA	PHOTOSYSTEM II REACTION CENTER PROTEIN A
PTM	envelope-bound plant homeodomain transcription factor
RDR6	RNA-DEPENDENT RNA POLYMERASE 6
RISC	RNA-induced silencing complex
RNA Pol	RNA polymerases
ROS	reactive oxygen species
SE	SERRATE
SGS3	SUPPRESSOR OF GENE SILENCING 3
SIG	sigma factor
siRNA	small interfering RNA
SMR	small MutS-related domain
SPL	SQUAMOSA PROMOTER BINDING PROTEIN LIKE
sRNAs	small non-coding RNAs
SRO5	SIMILAR TO RCD ONE 5
ssRNA	single-stranded RNA
ta-siRNA	trans-acting small interfering RNA
TE	transposable element
TIM	translocase of the inner membrane
TPB	tetrapyrrole biosynthesis pathway
TPR	tetratricopeptide repeat
XRNs	nuclear exoribonucleases

## **LIST OF PUBLICATIONS**

### **Publication I**

**Kristin Habermann**, Bhavika Tiwari, Maria Krantz, Stephan O. Adler, Edda Klipp, M. Asif Arif\*, Wolfgang Frank\*

**Impact of small RNAs in retrograde signalling pathways in *Arabidopsis thaliana***

<https://doi.org/10.1111/tpj.14912>

### **Publication II**

Bhavika Tiwari, **Kristin Habermann**, M. Asif Arif, Heinrich Lukas Weil, Timo Mühlhaus, Wolfgang Frank

**Identification of cold stress regulated small RNAs in *Arabidopsis thaliana***

<https://doi.org/10.1186/s12870-020-02511-3>

## **DECLARATION OF CONTRIBUTION AS A CO-AUTHOR**

### **Publication I (Habermann et al., 2020)**

Habermann, K. performed the research with help from Arif, M.A. and Tiwari, B.; Habermann, K., Arif, M.A., Tiwari, B and Frank, W. analysed the data; network analysis was performed by Adler, S.O., Krantz, M. and Klipp, E.; and Habermann, K., Arif, M.A., Adler, S.O., Krantz, M. and Frank, W. wrote the paper. All authors read and approved the final manuscript.

### **Publication II (Tiwari et al., 2020)**

Tiwari, B performed the research with help from Arif, M.A. and Habermann, K.; Tiwari, B, Arif, M.A., Habermann, K. and Frank, W. analysed the data; miRNA-TF network was constructed by Weil, H.L. and Mühlhaus, T.; and Tiwari, B, Arif, M.A., Weil, H.L. and Frank, W. wrote the paper. All authors read and approved the final manuscript.

We hereby confirm the above statements:

---

Kristin Habermann

---

Wolfgang Frank

## **SUMMARY**

Plants are sessile organisms that have to deal with environmental stress conditions due to changes in gene expression. The development of the chloroplast and the response to environmental fluctuations is dependent on signals released from plastid, which are mediated through retrograde signalling pathways causing alterations in gene expression. Further, abiotic stresses such as low temperatures negatively affect the plants and require alterations in gene expression to better cope with stress conditions. Posttranscriptional regulation through small non-coding RNA (sRNA) classes is one mechanism to regulate gene expression. However, deeper understandings of sRNAs and the regulation of their corresponding mRNA targets are still lacking. The aim of the thesis was to analyse sRNA, mRNA and long non-coding RNA (lncRNA) expression profiles in response to norflurazon in *Arabidopsis thaliana* wild type as well as the two retrograde signalling mutants, *gun1* and *gun5*. Further, sRNA expression profiling of the *Arabidopsis thaliana* wild type was performed in response to a time series (3 h, 6 h and 2 d) of cold acclimation condition (4°C). In both studies many different sRNA classes like microRNAs (miRNAs) have been detected to be differentially expressed in response to the applied treatments. Consequently, miRNA target prediction was performed to correlate the expression levels of miRNAs to their potential mRNA targets and to decipher their role in response to different stress conditions. In addition, for the miRNA:mRNA target pairs of interest, germination assays in miRNA overexpression lines and quantitative analysis of the putative targets have been performed to elucidate their impact in response to different stress conditions.

## **ZUSAMMENFASSUNG**

Pflanzen sind sessile Organismen, die sich aufgrund von Umweltstress mit Veränderungen der Genexpression auseinandersetzen müssen. Die Entwicklung des Chloroplasten und die Reaktion auf Umweltschwankungen hängen von plastidären Signalen ab, welche durch retrograde Signalwege vermittelt werden und Veränderungen in der Genexpression hervorrufen. Darüber hinaus beeinflussen abiotische Stressbedingungen wie niedrigere Temperaturen die Pflanzen negativ und fordern Änderungen in der Genexpression, um Stressbedingungen besser zu bewältigen. Die posttranskriptionelle Regulation durch kleine nicht-codierenden RNAs (sRNAs) ist eine Möglichkeit zur Regulierung der Genexpression. Allerdings fehlen immer noch tiefere Verständnisse über sRNAs und der Regulation ihrer korrespondierenden Ziel-mRNAs. Das Ziel der Dissertation war die Analyse von Expressionsprofilen von sRNAs, mRNAs und langen nicht-codierenden RNAs (lncRNAs) nach Norflurazon-Behandlung des *Arabidopsis thaliana* Wildtyps sowie zwei retrograden Signalmutanten, *gun1* und *gun5*. Weiterhin wurden sRNA-Expressionsprofile vom *Arabidopsis thaliana* Wildtyp unter Kälteakklimatisierung (4°C) in einer Zeitreihe (3 h, 6 h, 2 d) untersucht. In beiden Studien konnten nach abiotischen Stressbehandlungen eine Vielzahl an differenziell exprimierten sRNAs, wie microRNAs (miRNAs), ermittelt werden. Infolgedessen wurden Vorhersagen zu möglichen miRNA Ziel-mRNAs durchgeführt, um die Expressionsniveaus der miRNAs mit der Regulierung ihrer Ziel-mRNAs zu korrelieren, um ihre Rolle als Reaktion auf unterschiedliche Stressbedingungen zu entschlüsseln. Zusätzlich wurden für vielversprechende miRNA:RNA Paare Keimungsassays mit verschiedenen miRNA-Überexpressionslinien und eine quantitative Analyse der Ziel-mRNAs durchgeführt, um deren Auswirkungen in Reaktion auf verschiedene Stressbedingungen zu erklären.

## **INTRODUCTION**

### **Endosymbiotic theory**

More than 100 years ago, in 1905, Mereschkowsky published a paper suggesting that organelles contain genes (Martin and Kowallik, 1999). However, it took many more years until it was partially accepted that mitochondria and chloroplasts have evolved through the endosymbiosis of distinct prokaryotic progenitors, an evolutionary event known as endosymbiotic theory today (Margulis, 1970). Whereas the endosymbiosis of  $\alpha$ -proteobacteria led to the formation of mitochondria in all common eukaryotes, plastids, present only in the green lineage of life, have evolved from endosymbiotic cyanobacteria; however, until now it is unknown which lineages exactly give rise to the specific organelles (Gray et al., 1999, Martin et al., 2002, Wu et al., 2004). During the functional internalisation and transformation into plastids, the major part of the cyanobacterial genome was transferred to the nuclear DNA of the host organism, and only few protein coding genes remained in the plastid genome (Simpson and Stern, 2002). The number of protein coding genes, which have been transferred, were specific to the donor cyanobacteria species as well as acceptor eukaryotes (Timmis et al., 2004). Hence, several genes necessary for organellar development and regulatory functions are now located in the nuclear genome and their proteins have to be imported into the plastid (Martin et al., 2002). As a result of this gene transfer, the stoichiometry of nuclear and chloroplastic encoded polypeptides functioning in the plastids requires the coordination of both genomes during various developmental stages as well as in response to environmental stimuli (Zimorski et al., 2014).

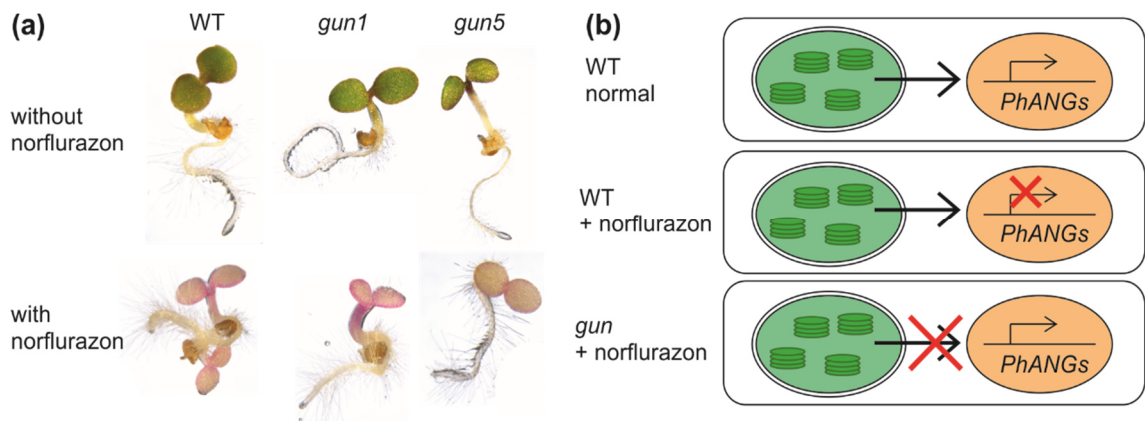


### **Anterograde and retrograde signalling**

In plants, there are two pathways that highly coordinate the nuclear and organellar gene expression. Firstly, there is the anterograde signalling, which controls organellar gene expression by nucleus-to-plastid signals. It is required for the direct regulation of the organellar gene expression through different nuclear-encoded factors that are transported into the plastid. The anterograde signalling modulates chloroplast transcription, RNA metabolism, and the post-transcriptional and post-translational regulation through RNA-binding proteins, such as pentatricopeptide repeat (PPR) and tetratricopeptide repeat (TPR) proteins, respectively (Berry et al., 2013). The second pathway is the so-called retrograde signalling, which controls nuclear genes through plastid-to-nucleus signals (Kleine and Leister, 2016). In contrast to the anterograde signalling, the retrograde signalling utilizes signals, which are sent from the chloroplast (or mitochondria) to the nucleus, and thus, affected the expression of nuclear-encoded genes (Surpin and Chory, 1997). These factors are known to originate from (i) plastid gene expression, (ii) plastid metabolites like the tetrapyrroles, (iii) the reactive oxygen species (ROS) and changes in the redox state (Kleine et al., 2009, Terry and Smith, 2013, Dietz et al., 2016, Kim, 2019, Zhao et al., 2019a). For instance, lincomycin, a plastid translation inhibitor, or rifampicin, a selective inhibitor of the plastid-encoded RNA polymerase (RNA Pol), cause defects in gene expression and trigger retrograde signalling, and thus lead to a decrease of the photosynthesis associated nuclear genes (*PhANGs*) (Sullivan and Gray, 1999). Another example is norflurazon (NF), a herbicide that acts as a non-competitive inhibitor of the phytoene desaturase enzyme (Oelmüller and Mohr, 1986). The phytoene desaturase catalyses a two-step dehydrogenation of phytoene to generate zeta-carotene in the carotenoid biosynthesis pathway (Suarez et al., 2014). The carotenoids as natural antioxidants are components of the light-harvesting complexes that are involved also in chloroplast protection from negative effects of ROS through a mechanism known as nonphotochemical quenching (Kim and Apel, 2013). An oxidative by-product of  $\beta$ -carotene is  $\beta$ -cyclocitral, known to be a retrograde signal (Ramel et al., 2012). Therefore, NF blocks the carotenoid biosynthesis pathway by inhibiting the phytoene

desaturase, resulting in photooxidative damage and bleaching of the green plant tissue (Figure 1a); moreover, the absence of  $\beta$ -cyclocitral reduces expression of *PhANGs* (D'Alessandro et al., 2018).

To gain more knowledge about the factors involved in retrograde signalling during chloroplast development, several mutant screens were performed in *Arabidopsis thaliana* (Susek et al., 1993, Larkin et al., 2003, Mochizuki et al., 2001, Gray et al., 2003, Saini et al., 2011, Meskauskienė et al., 2001, Gutierrez-Nava et al., 2004, Ball et al., 2004, Rossel et al., 2006, Woodson et al., 2011). Transgenic wild type plants containing the  $\beta$ -glucuronidase (*GUS*) reporter gene and the light harvesting complex of photosystem II (*LHCB*) promoter were mutagenized with ethyl methanesulfonate (EMS), and subsequently screened in presence of NF with *GUS* activity and the mutants show blue cotyledons (Susek et al., 1993). Further, wild type and selected mutagenized seedlings were grown in the presence of NF to monitor the expression of some *PhANGs* like *LHCB1.2* (Figure 1b). Thereby, GENOME UNCOUPLED (*gun*) mutants were isolated, which showed a de-repression of *PhANGs* despite chloroplast development perturbations induced by NF (Susek et al., 1993, Woodson et al., 2011).



**Figure 1: Retrograde signalling mutants.** Images of four-day-old seedlings of the wild type, *gun1* and *gun5* mutants of *Arabidopsis thaliana* treated with and without NF (a) (adapted from (Habermann et al., 2020)). The retrograde signalling can result in following reduced pattern (b). Under normal conditions in wild type the *PhANGs* can be expressed normally, but when treated with NF a repressor signal can move from chloroplast to nucleus that inhibits the gene expression of the *PhANGs*; in case of *gun* mutants the retrograde signalling pathway is disturbed and the absence of a signalling molecule results in a gene expression even under photobleaching conditions.

Five of the six *gun* mutants, namely *gun2* to *gun6*, were found to be involved in the tetrapyrrole biosynthesis pathway (TPB). *GUN2* encodes a heme oxygenase, which converts heme to biliverdin IX $\alpha$ . *GUN3* encodes a phytochromobilin synthase, which is required for the generation of phytochromobilin through biliverdin IX $\alpha$  (Mochizuki et al., 2001). Both *gun2* and *gun3* mutants are affected in the heme biogenesis that provoke a higher accumulation of heme in the plastid. Moreover, studies have shown that the *gun2* is allelic to the long hypocotyl 1 (*hy1*) mutant and *gun3* is allelic to the *hy2* mutant, which leads to the assumption that *hy1-1* and *hy2-1* are *gun* mutants (Koornneef and van der Veen, 1980, Mochizuki et al., 2001). The *GUN4* gene encodes a protoporphyrin IX (Proto)/ magnesium-protoporphyrin IX (Mg-Proto) binding protein supporting the accumulation of chlorophyll by leading to a higher activation of the magnesium-chelatase, which means that the *gun4* mutant has an increased efficiency of this enzyme (Larkin et al., 2003, Adhikari et al., 2011). The *GUN5* gene encodes the CHLH subunit of the magnesium-chelatase, which catalyses the conversion of Proto to Mg-Proto in the chlorophyll biosynthetic pathway. The *gun5* mutant is defective in the *CHLH* gene due to a nucleotide substitution (C to T) and produces a less efficient magnesium-chelatase (Mochizuki et al., 2001). Moreover, the *gun6* mutant was identified to have an increased ferrochelatase I expression, which resulted in higher ferrochelatase activity (Woodson et al., 2011, Woodson et al., 2013). In summary, these studies showed the involvement of tetrapyrroles in retrograde signalling, but how these compounds exactly contribute to plastid-to-nucleus signalling is still not clear (Mochizuki et al., 2001, Mochizuki et al., 2008, Strand et al., 2003). Finally, the *gun1* mutant differs from the other *gun* mutants, because it is able to express *PhANGs* not only after exposure to NF, but also in the presence of lincomycin (Shimizu et al., 2019). *GUN1* encodes a chloroplast-localised pentatricopeptide repeat (PPR) protein containing a C-terminal small MutS-related domain (SMR) (Tadini et al., 2016, Barkan and Small, 2014). Even though the exact molecular mechanism of *GUN1* is still unknown, it is suggested that the *gun1* mutant can integrate signals derived from perturbations in the TPB, plastid gene expression and redox state (Kleine and Leister, 2016). Previous genetic evidences suggested that *GUN1* signalling activates the nuclear

transcription factor *ABI4* through a chloroplast envelope-bound plant homeodomain transcription factor (*PTM*) (Koussevitzky et al., 2007, Sun et al., 2011). However, two recent studies could not confirm that *PTM* or *ABI4* play a role in the *GUN1*-mediated retrograde signalling pathway, since no consistent *gun* phenotype could be detected under various conditions in both mutants (Page et al., 2017, Kacprzak et al., 2019).

However, *GUN1* plays an important role during the regulation of the chloroplast development in response to different changes disturbing protein homeostasis (Llamas et al., 2017). Recently, Wu *et al.* (2018) discovered that *GUN1* acts in two main networks, the so-called biogenic and operational control. During the development of chloroplast and photosystem the biogenic control takes place and the operational control acts through fluctuations in the environment in later stages (Kleine and Leister, 2016). In conclusion, more investigations are still required to reveal all the biological functions of *GUN1* (Pesaresi and Kim, 2019).

The *GOLDEN2-LIKE* (*GLK*) transcription factors are involved in expression of nuclear photosynthetic genes and development of chloroplast. Interestingly, the double mutant *glk1glk2* was identified to exhibit a subtle *gun* phenotype with reduced chlorophyll intermediates that can affect the *GUN*-mediated retrograde signalling pathway (Waters et al., 2009, Leister and Kleine, 2016). In addition, the overexpression lines of *GLK1* and *GLK2* show higher expression levels of *LHCB* transcripts (Waters et al., 2009). The overexpression lines of *GLK1* and *GLK2* can be compared phenotypically to the *gun1-1* and *gun5-1* mutants in response to NF, which can make a link to the organellar gene expression (Leister and Kleine, 2016).

In the last years, additional studies and screenings were performed to shed more light on the retrograde signalling pathway. For instance, the retrograde signalling pathway was also found to be involved in the photoreceptor pathways in regulating photomorphogenesis (Martin et al., 2016), the control of the flowering time (Feng et al., 2016), acclimation in response to abiotic stress (Leister et al., 2017) as well as hormonal signalling cascades (Bobik and Burch-Smith, 2015, Gollan et al., 2015).

Thus far, two microarray studies investigated the impact of plastid-to-nucleus signals by comparing transcriptional changes in the *Arabidopsis thaliana* *gun* mutants and wild type (Strand et al., 2003, Koussevitzky et al., 2007). The first transcriptome study was performed by Strand *et al.* (2003) and compared wild type, *gun1*, *gun2* and *gun5* mutants after growing for six days in the presence of NF (Strand et al., 2003). The second transcriptome study focused on wild type together with *gun1* and *gun5* mutants after five days of NF treatment (Koussevitzky et al., 2007). Both studies could detect a robust link between the retrograde signalling mutants *gun1* and *gun5* (and *gun2* in case of Strand *et al.* (2003)), because a considerable number of detected genes had equal expression patterns. Additionally, it was identified that *PhANGs* are mainly decreased in the wild type samples; however these were de-repressed in the *gun* mutants in response to NF.

**Biogenesis and functions of small non-coding RNAs**

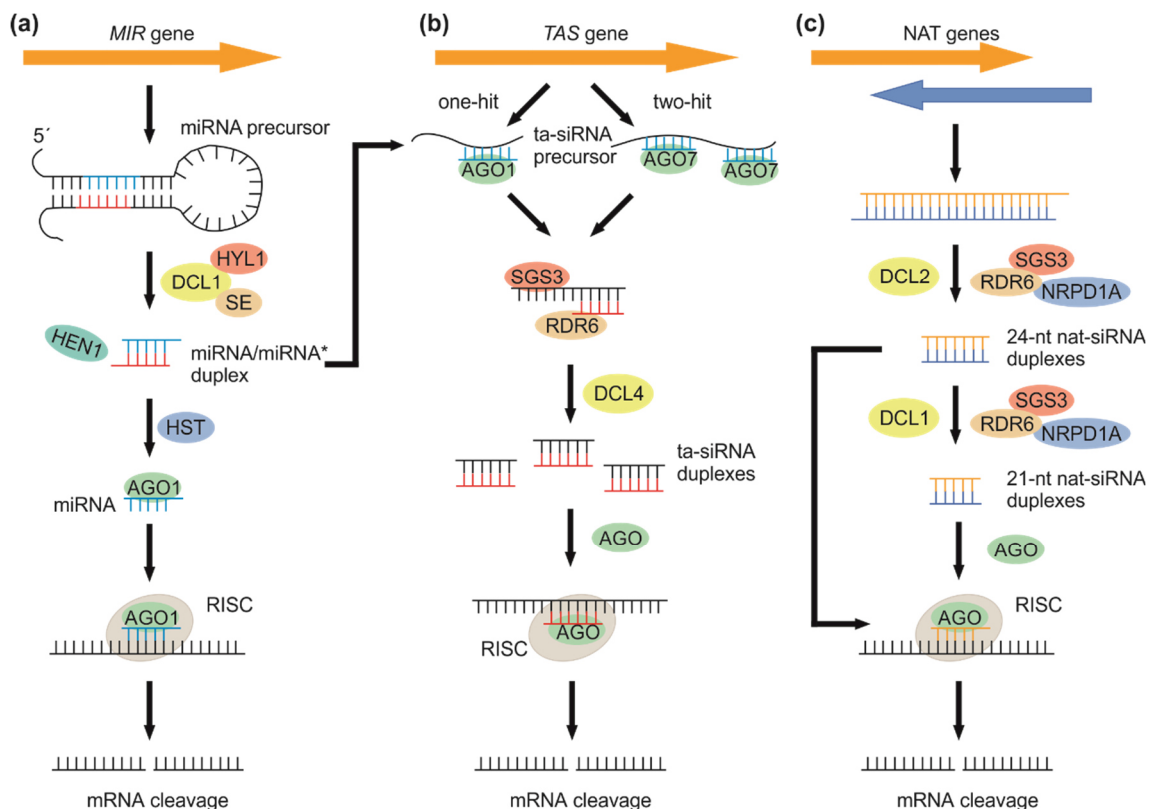
So far, nearly all studies have used transcriptome analysis to investigate the alteration of protein-coding gene expression in retrograde signalling mutants (Strand et al., 2003, Koussevitzky et al., 2007, Marino et al., 2019, Richter et al., 2020). Indeed, it is well known that there are also other regulatory RNA classes that have an impact on several biological plant functions. The class of non-coding RNAs (ncRNAs) including long ncRNAs (lncRNAs) and small ncRNAs (sRNAs) are known to play important roles during plant development and in adaptation to various abiotic and biotic stresses, as the ncRNAs are involved mostly in gene expression regulation (Wang and Chekanova, 2017, Huang et al., 2019).

The class ncRNAs of referred as sRNAs are 20 - 24 nucleotides in length and control gene expression in many biological processes (Li et al., 2017). For instance, sRNAs are able to control epigenetic modifications by transcriptional gene silencing (Holoach and Moazed, 2015, Bannister and Kouzarides, 2011, Khraiweh et al., 2010). However, it is more common that sRNAs mediate the cleavage or translational inhibition through binding to the mRNA target, and thus function in the post-transcriptional control (Catalanotto et al., 2016). sRNAs can be classified according to their origin, since they can originate from a double-stranded RNA (dsRNA) or a single-stranded RNA (ssRNA) precursor (Axtell, 2013). Those sRNAs processed from the dsRNA precursor are called small interfering RNA (siRNA), but sRNAs can also be processed from a ssRNA precursor named hairpin RNA (hpRNA).

One class of hpRNA generating microRNAs (miRNAs), is well-studied and known to control transcript levels by sequence specific base pairing during various biological processes (Reinhart et al., 2002, Manavella et al., 2019). The RNA Pol II transcribes the endogenous *MIR* genes in ssRNAs, which form a characteristic hairpin structure due to intra-molecular sequence complementarity. This results in a 5' capped and polyA-tailed primary miRNA (pri-miRNA) transcript. Further, the pri-miRNA has to be cleaved into precursor miRNA (pre-miRNA) through two cleavages, whereas the first cleavage determines the mature miRNA and the second cleavage is placed at the end of the *MIR* precursor. This cleavage occurs

through DICER-LIKE1 (DCL1) enzymes with the help of HYPONASTIC LEAVES 1 (HYL1) and SERRATE (SE) and they form the so-called dicing body (Laubinger et al., 2008). In the next step, the pre-miRNA is further processed into a miRNA/miRNA\* duplex with 20 - 22 nucleotide in length by DCL1, and the 3' ends of both strands of the miRNA/miRNA\* duplex are 2'-O-methylated by HUA ENHANCER 1 (HEN1), protecting those strands from degradation (Yu et al., 2005, Park et al., 2002). The miRNA/miRNA\* duplex comprises the mature miRNA (guide strand) as well as the miRNA\* (miRNA passenger strand). The place of function of mature miRNAs is the cytosol, but the miRNA/miRNA\* duplex is still located in the nucleus, which requires now to be exported by RNA exporters HASTY (HST) (Park et al., 2002). The ARGONAUTE1 (AGO1) protein acts as a central module in the miRNA biogenesis and function, which loads the mature miRNA strand from the miRNA/miRNA\* duplex to form the RNA-induced silencing complex (RISC) (Rogers and Chen, 2013). The miRNA\* strand is usually degraded, but in some cases the strand can be stabilised und functionalised during stress conditions (Devers et al., 2011). The RISC guides the miRNA to its mRNA target by Watson-Crick pairing, leading to a successive transcript cleavage at position 10 or 11 and the mRNA gets degraded by nuclear exoribonucleases (XRN) or the RNA target acts a precursor for the biogenesis of other siRNAs such as trans-acting siRNA (ta-siRNA) (Figure 2a) (Wierzbicki et al., 2008, Voinnet, 2009). With those functions, miRNAs are able to mediate the regulation of the transcript levels of many transcription factors or stress responsive proteins under various biological processes, such as plant development, signal transduction or in response to environmental stress conditions (Yu et al., 2017). Moreover, in a recent study by Fang *et al.* (2018), it has been detected that sRNAs can have an impact on retrograde signalling, since they discovered that tocopheroles and 3'-phosphoadenosine 5'-phosphate (PAP) have an effect on miRNA processing. The metabolite PAP was previously identified to have a function in the RNA metabolism, where it serves as a mobile signal to inhibit XRN, and thus preventing the degradation of 5' uncapped ends of mRNA and pri-miRNA (Estavillo et al., 2011). Fang *et al.* (2018) showed that tocopheroles are positive regulators of miRNA biogenesis, since they also

positively control PAP accumulation. XRN2 levels can be inhibited by PAP, leading to an increased pri-miRNA accumulation (Fang et al., 2018). Furthermore, miR395 is able to guide cleavage of a transcript encoding the enzyme ATP sulfurylase (APS), which catalyses the initial step of the PAP synthesis, and miR398 targets the mRNA of *COPPER/ZINC SUPEROXIDE DISMUTASE 2* (*CSD2*) for cleavage under heat stress, causing not only a negative regulation of miRNA biogenesis through a reduction of the synthesis of PAP precursors, but also protection against heat through *CSD2* regulation (Fang et al., 2018).



**Figure 2: Biogenesis pathway of different sRNAs.** Different sRNAs are processed through different pathways. (a) The biogenesis of miRNAs comes along with a hairpin structure of the miRNA precursor and those cleavage is mediated through DCL1. (b) The *TAS* precursor gets cleaved through AGO proteins with the help of miRNAs and the ta-siRNA processing occurs through dsRNA synthesis requiring the DCL4 phased cleavage. (c) nat-siRNAs are processed through the overlap of two mRNA transcripts and subsequent cleavages by DCL2 and DCL1.

In addition to hpRNAs, the other class of siRNAs that can be processed from dsRNA precursors include natural antisense transcript-derived siRNA (nat-



siRNA) and trans-acting siRNA (ta-siRNA). ta-siRNAs are generated with the help of miRNAs, which target and guide cleavage of the *TAS* precursor transcript, leading to a dsRNA by RNA-DEPENDENT RNA POLYMERASE 6 (RDR6), which is subsequently processed to phased siRNAs by DCL4. In *Arabidopsis thaliana*, four different *TAS* genes, encoded by eight different loci, were identified *TAS1a/b/c*, *TAS2*, *TAS3a/b/c* and *TAS4* (Allen et al., 2005, Rajagopalan et al., 2006). Two different pathways are known to produce ta-siRNAs: the one-hit and two-hit pathway. The precursors of *TAS1*, *TAS2* and *TAS4* possess just one single miRNA binding site and so they are part of the one-hit pathway, whereas the remaining *TAS3* precursor has two miRNA binding sites, and thus belongs to the two-hit pathway (Yoshikawa, 2013). *TAS1* and *TAS2* transcripts are targeted by miR173, which has a size of 22 nucleotides, and the *TAS4* transcript is targeted by miR828 with a length of 22 nucleotides. In the one-hit pathway, the cleavage of the *TAS* mRNA precursor is mediated through AGO1 and miR173 or miR828. On the other hand, AGO7 and the miR390 with 21 nucleotides in size mediate the cleavage of the *TAS3* transcript at two diverse binding sites (Singh et al., 2018). Nevertheless, after cleaving of the *TAS* transcripts, SUPPRESSOR OF GENE SILENCING 3 (SGS3) is relevant for stabilising the ssRNAs, RDR6 generates dsRNAs and DCL4 cleaves those dsRNAs into phased ta-siRNAs with 24 nucleotides in size that are subsequently loaded in the RISC complex to regulate their RNA targets (Figure 2b) (Borges and Martienssen, 2015). Different ta-siRNAs are known to have different targets. Those ta-siRNAs which are processed from *TAS1* and *TAS2* precursor target PPR proteins (Axtell et al., 2006). In addition, ta-siRNAs originated from the *TAS2* precursor are mainly targeting transcripts encoding *AUXIN RESPONSE FACTORS* (*ARF*), and thus control the auxin signalling pathway (Allen et al., 2005, Fahlgren et al., 2006). Transcripts of *MYB* transcription factors or *PRODUCTION OF ANTHOCYANIN PIGMENT 1* (*PAP1*) are targets of the ta-siRNAs derived from the *TAS4* precursor to regulate the anthocyanin biosynthesis (Luo et al., 2012).

In contrast to ta-siRNAs, dsRNA molecules underlying nat-siRNA biogenesis are generated through the overlapping of two endogenous transcripts in antisense orientation. Originally, the biogenesis of nat-siRNAs was discovered in salt stress

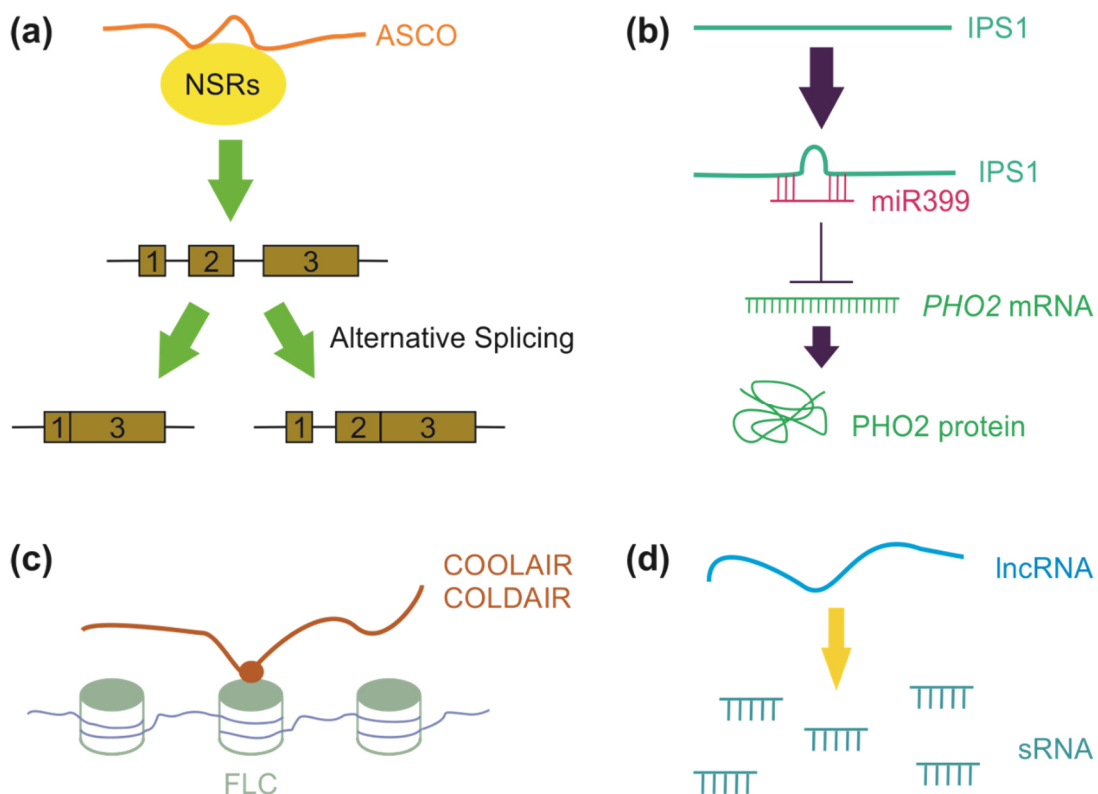
experiment, where nat-siRNAs were processed from two overlapping transcripts (Borsani et al., 2005). In particular, ROS accumulation triggered by salt stress causes a higher expression of *SIMILAR TO RCD ONE 5 (SRO5)*. Borsani et al. (2005) detected an overlapping of the mRNA transcripts of *SRO5* and *DELTA1-PYRROLINE-5-CARBOXYLATE DEHYDROGENASE (P5CDH)* encoding for a constitutively expressed enzyme that catabolises the production of DELTA1-PYRROLINE-5-CARBOXYLATE (P5C) in the proline metabolism. DCL2 initiates the cleavage of the overlapping region of the mRNA transcripts of *SRO5* and *P5CDH*, leading to the formation of double stranded nat-siRNAs with a size of 24 nucleotides. This step involves also NUCLEAR RNA POLYMERASE D1A (NRPD1A), RDR6 and SGS3. One strand of the 24-nucleotide double stranded nat-siRNA has two functions. The first one is to target the *P5CDH* mRNA transcript for cleavage, and the second function is to phase the successive cleavage by DCL1, NRPD1A, RDR6 and SGS3 that generates a formation of nat-siRNAs. Those 21-nucleotide nat-siRNAs also provoke the cleavage of the *P5CDH* mRNA transcript (Figure 2c). Moreover, reduced expression levels of mRNA transcript of *P5CDH* cause less proteins, which in turn leads to an accumulation of proline, P5C and ROS that are involved in salt stress adaptation mechanisms (Borsani et al., 2005).

Within the nat-siRNA class, two different types can be distinguished based on the chromosomal arrangement of the overlapping endogenous transcripts: *cis*- and *trans*-nat-siRNAs (Yuan et al., 2015). The nat-siRNA are called *cis*-nat-siRNA when the RNA of the overlapping genes is transcribed from opposite DNA strands in an equal genomic region, whereas, if both genes encode RNA transcripts from different genomic regions and have reverse complementary sequences, then they are able to generate *trans*-nat-siRNA (Lapidot and Pilpel, 2006). Regardless of the chromosomal arrangement, both nat-siRNA types have the same function: they are able to cause the cleavage of one of the initial overlapping mRNA transcripts, thus leading to gene expression regulation in response to various biological mechanisms (Borsani et al., 2005, Katiyar-Agarwal et al., 2006, Zhang et al., 2012).

There are many other sRNAs known, which do not belong to any of the already described classes, because of their different biogenesis pathways, for instance phased small interfering RNAs (phasiRNAs), heterochromatic siRNAs (hetsiRNAs), epigenetically-activated small interfering RNAs (easiRNAs) or piwi-interacting RNA (piRNAs) (Axtell, 2013, Borges and Martienssen, 2015, Singh et al., 2018).

### Biogenesis and functions of long non-coding RNAs

The ncRNAs possessing a size longer than 200 nucleotides are called lncRNAs. Generally, lncRNAs are transcribed by different RNA polymerases (RNA Pol II, III, IV or V), but do not produce a functional protein (Wierzbicki et al., 2008). They are known to regulate gene expression in different biological processes, such as fertility, photomorphogenesis, protein re-localization or modulation of chromatin loop dynamics (Liu et al., 2015).



**Figure 3: lncRNAs and their functions.** lncRNAs can have various functions in regulating different biological mechanisms. The lncRNA ASCO is known to cause alternative splicing events (a). The lncRNA IPS1 acts through target mimicry to serve as a decoy for miR399 to increase the expression levels of the miRNA target, *PHO2*, to maintain the inorganic phosphate (Pi) homeostasis (b). In addition, lncRNAs like COOLAIR and COLDAIR guide histone modifications under cold temperatures at the *FLC* locus to cause a delay in the flowering time (c). Furthermore, lncRNAs can act as precursors of other siRNAs (d).

One important role is the regulation of alternative splicing events through activation or inhibition of mRNA splicing (Figure 3a). It was discovered that the alternative splicing competitor (ASCO) lncRNA interacts with the nuclear speckle RNA-binding protein (NSR) to build an alternative splicing regulatory module,

which controls lateral root development (Bardou et al., 2014). Another interesting function of lncRNAs is target mimicry, where the lncRNA acts as a decoy for miRNAs by complementing and sequestering miRNAs (Figure 3b). One well known biological example in *Arabidopsis thaliana* is the endogenous lncRNA INDUCED BY PHOSPHATE STARVATION1 (IPS1), which is increased during phosphate excess and operates as a target of miR399 (Franco-Zorrilla et al., 2007). Typically, the *UBIQUITIN-CONJUGATING ENZYME 24 (PHO2)* is an important enzyme to maintain the inorganic phosphate (Pi) homeostasis, as it ubiquitinates proteins for further degradation such as the Pi transporters (*PHT1*) that need to be reduced under high Pi (Aung et al., 2006). The mRNA transcript of the *PHO2* is a target of miR399, which regulates *PHO2* protein levels to further regulate the downstream Pi uptake (Chiou et al., 2006). The lncRNA IPS1 acts as binding partner and sequesters miR399 with a mismatch creating loop around miRNA cleavage site leading to increased *PHO2* levels and enhanced ubiquitination of cognate protein targets and subsequent degradation resulting in a survival of plants under high Pi (Franco-Zorrilla et al., 2007). Another interesting biological function of lncRNAs is the control of vernalization that promotes flowering time through epigenetic modifications (Figure 3c). Commonly, the *FLOWERING LOCUS C (FLC)* plays a key role in *Arabidopsis thaliana* to delay flowering under low temperatures (Michaelis and Amasino, 1999). There are two lncRNAs transcribed from the *FLC* locus: COLD INDUCED LONG ANTISENSE INTRAGENIC RNA (COOLAIR) that is transcribed in antisense direction (Swiezewski et al., 2009), and COLD ASSISTED INTRONIC NONCODING RNA (COLDAIR) that is transcribed in the sense direction (Heo and Sung, 2011). Both lncRNAs, COOLAIR and COLDAIR, act as epigenetic regulators and cause histone modifications on the *FLC* locus, thus inhibiting flowering under cold temperatures (Heo and Sung, 2011). Interestingly, lncRNAs can also serve as precursors for sRNAs (Figure 3d) (Ma et al., 2014). Tang *et al.* (2019) discovered numerous RDR2- and DCL3-dependent sRNAs *in silico*, which have been generated from lncRNAs in *Arabidopsis thaliana* and the sRNAs have been detected to be associated with AGO4 for chromatin modifications or associated with AGO1 for target cleavage, but further validations are needed.

### **Non-coding RNAs in biotic and abiotic stress response in *Arabidopsis thaliana***

Plants have developed complex regulatory mechanisms to adapt to stressful conditions like variations in climate and environmental conditions (abiotic stresses) or upon attack by a vast range of pests and pathogens (biotic stresses). One important regulatory mechanism involves the regulation of mRNA transcript levels of transcription factors and stress responsive proteins via RNA interference (RNAi) mediated by sRNA (Yang et al., 2018b, Ben-Gera et al., 2016).

Upon pathogen attack, the plant immune response is mediated by the so-called resistance genes, which comprise the most common nucleotide-binding leucine-rich repeat (NB-LRR) proteins (Ellis et al., 2000). The activated resistance genes induce also re-allocation of plant resources; indeed, an autoimmune response triggered by unregulated resistance gene expression is able to restrain plant growth (Zhai et al., 2011, Li et al., 2012). *Pseudomonas syringae* is a ubiquitous gram-negative bacterium that causes many different plant diseases such as bacterial canker and makes plants more fragile for environmental changes like low temperatures. In *Arabidopsis thaliana*, during *Pseudomonas syringae* infection, a flagellin-derived peptide triggers the induction of miR393, which targets mRNA transcripts of auxin receptors, and thus suppresses the auxin signalling and shifts the plant priority from development towards defence (Navarro et al., 2006). Further investigations showed that miR393 also affects the salicylic acid pathway, since it inhibits auxin signalling, which causes decreased levels of camalexin, an indolic phytoalexin preventing necrotrophic fungi, and higher levels of glucosinolate, a natural component preventing pests and diseases, that boost *Arabidopsis thaliana* immunity to *Pseudomonas syringae* (Robert-Seilantantz et al., 2011). Another miRNA that was identified to play a role during *Pseudomonas syringae* infection is miR863-3p, which targets the mRNA transcripts of atypical receptor-like pseudokinases, *ARLPK1* and *ARLPK2* (Niu et al., 2016). Interestingly, the mRNA cleavage of *ARLPK1* and *ARLPK2* directed by miR863-3p is an early defence response upon infection. The transcript of *SE* provides another target for miR863-3p, and the transcriptional inhibition of *SE* causes a self-regulation of miR863-3p (Niu et al., 2016).

Different stress conditions, such as oxidative stress, hypoxia, extreme temperatures, high salinity and drought, are able to trigger RNAi-based gene silencing mechanisms mediated by sRNAs, in particular by miRNAs (Khraiwesh et al., 2012, Li et al., 2017, Shriram et al., 2016). Moreover, miRNAs have been identified to play not only a role under one stress condition, but they are also able to be affected by many stress conditions like the highly conserved miR408 and miR156 (Ma et al., 2015, Zheng et al., 2019).

Plant hormones, especially auxin, are main focus of studies as they control developmental processes including root hair elongation. In *Arabidopsis thaliana*, the ARF transcription factors are known to play a role during auxin signal transduction in plant development; and the mRNA transcripts of *ARF10* are targeted by miR160 (Liu et al., 2007). Furthermore, overexpression lines of miR160 have been described to have reduced sensitivity against abscisic acid (ABA) during seed germination (Liu et al., 2007). Further, miR160 has an important role during seed germination as well as post-germination under stress conditions (Liu et al., 2007).

Several miRNAs were identified to play a role mainly during drought and salt stress; as most of the studies were focused on those conditions (Khraiwesh et al., 2012, Shriram et al., 2016). The miRNA miR394 was identified to be induced by ABA; and transgenic lines overexpressing miR394 are very sensitive to high salt concentration, but more tolerant to drought stress (Song et al., 2013). In *Arabidopsis thaliana*, the *LEAF CURLING RESPONSIVENESS (LCR)* is essential to maintain the leaf morphology and its mRNA transcript serves as a target for miR394, and this interaction is required for a proper leaf shape, which can help to adapt to stress conditions like salt or drought (Song et al., 2012). Another example is miR399f, involved in phosphate homeostasis that was also found to be implicated in stress response related to ABA, different salt concentrations and drought (Baek et al., 2016). In this context, one already validated miRNA target transcript encoding for MYB DOMAIN PROTEIN 2 (AtMYB2) and two additional putative miRNA targets encoding for *ABA-RESPONSIVE ELEMENT (ABRE)-BINDING TRANSCRIPTION FACTOR3 (ABF3)* and *CHLOROPLAST STEM-LOOP BINDING PROTEIN OF 41 KDA*

(*CSP41b*) were suggested to play a role under those stress conditions. Overexpression lines of miR399f were found to increase an ABA and salt tolerance along with reduced root growth and seed germination as well as a hypersensitivity to drought stress (Baek et al., 2016). Therefore, it was concluded that miR399f plays a role in adaptive response to abiotic stress conditions by regulating its targets, *ABF3* and *CSP41b*, to maintain the balance of the homeostasis associated with ABA, different salt concentrations and drought.

Different stresses influence the plants in terms of development and productivity. For example, the population is increasing, which result in a future problem to feed the world population and the focus of studies should be on crops to increase their efficiency since particularly for instance under low temperatures (Chinnusamy et al., 2007, Ding et al., 2019, Kazemi-Shahandashti and Maali-Amiri, 2018). Cold stress can be categorised according to the temperature: freezing stress when the temperatures are below 0°C, and chilling stress when the temperatures are below 20°C. At low temperatures, two pathways are responsible for acclimation in plants: the C-repeat/drought-responsive element binding factor (CBF)-dependent and CBF-independent transcriptional pathway (Zhu, 2016). The inducer of CBF expression 1 (*ICE1*) regulates the expression of the genes coding for CBF transcription factors by binding on their promoter region (Chinnusamy et al., 2010). Further, *ICE1* is also involved in the activation of the expression of dehydration responsive element binding (DREB) transcription factors, which in turn bind to the cold response sensitive transcription factors/dehydration responsive elements (CRT/DRE) promotor regions of the cold-responsive (*COR*) genes that encode cold responsive proteins. Another CBF-independent transcriptional pathway regulates the expression of *COR* genes, which is triggered by the binding of basic leucine zipper (bZIP) transcription factors ABA-responsive element (ABRE)-binding protein (AREB)/ABRE-binding factors (ABFs) and those AREB/ABF proteins verified an ABA-mediated signal (Huang et al., 2012). Lee *et al.* (2010) discovered a direct interaction between ABA-mediated stress response due to ABF transcription factors and DREB/CBF gene family, which suggests a crosstalk between different stress responsive pathways.



Several studies have shown that sRNAs, especially miRNAs, play an important role in regulating transcription factor and *COR* genes under cold stress (Pegler et al., 2019, Sunkar and Zhu, 2004, Liu et al., 2008, Megha et al., 2018). Overexpression lines of miR397a in *Arabidopsis thaliana* were investigated using cold tolerance assay, implying an improved adaptation of those overexpression lines at low temperatures (Dong, 2014). In particular, the expression of the mRNA transcripts of the *CBF* genes and their downstream *COR* genes are affected under cold treatment in miR397a overexpression lines, leading to the conclusion that miR397a plays a role in controlling the cold signalling pathway (Dong, 2014). In *Arabidopsis thaliana*, leaf senescence is regulated by miR408, which guides the cleavage of mRNA transcripts of a cupredoxin superfamily protein, *LACCASE 3* and *PLANTACYANIN* that are expressed in leaves (Thatcher et al., 2015). Those mRNA transcripts regulated by miR408 belonging to the group of cupredoxins acting as electron transfer shuttles between proteins in several biological processes (Choi and Davidson, 2011). The copper containing protein laccase (*LACCASE 13*) is implicated in the polymerization of lignin and miR408 overexpression lines resulted in decreased mRNA *LACCASE 13* expression levels and increased seed yield and biomass (Song et al., 2017). Many studies focus on the highly conserved miR408 and their targets, because of their important functions in regulation within different biological pathways in response to various stress conditions like copper, high salt or drought in different plant species for instance *Arabidopsis thaliana*, rice, maize (Liu et al., 2008, Rajwanshi et al., 2014, Trindade et al., 2010, Hajyzadeh et al., 2015, Zhang et al., 2017). Flowering time is mainly controlled by members of the SQUAMOSA PROMOTER BINDING PROTEIN LIKE (*SPL*) transcription factors and their mRNA transcripts function as targets for the miRNA family miR156 (Wu et al., 2009). Under abiotic stress conditions, miR156 targets the *SPL* mRNA transcript, causing a delay flowering and thus keeping the plant in a juvenile state (Cui et al., 2014, Stief et al., 2014).

## **AIM OF THE THESIS**

Until now, only less sequencing studies focused on the correlation between sRNAs and mRNA transcripts (Crisp et al., 2017, Thatcher et al., 2015, Vidal et al., 2013). The function of sRNAs and other ncRNAs has not fully studied in the retrograde signalling pathway. Therefore, we wanted to make use of next generation sequencing (NGS) to investigate the role of sRNAs and other ncRNAs in two well-known retrograde signalling mutants, *gun1* and *gun5*. We intended *Arabidopsis thaliana* wild type, *gun1* and *gun5* seeds in the presence/absence of NF, and whole RNA, including sRNA, ncRNA and mRNA, were isolated from four-day-old seedlings and sequenced. All RNA profiles were analysed in terms of changes in expression levels and correlated to the specific retrograde signalling pathway. Consequently, the sRNA profiles have been associated with lncRNA and mRNA expression levels to determine how sRNAs are involved in the retrograde signalling pathway.

Additionally, the ncRNAs were also investigated in two-weeks-old *Arabidopsis thaliana* wild type plants in response to cold treatment. Plants growing at 4°C were collected at different time points, 3 h, 6 h, and 2 d, and different expression levels of sRNA have been correlated to ncRNA and mRNA from a public data set in order to gain more knowledge about cold responses mediated by sRNA or ncRNA (Garcia-Molina et al., 2020).

Preliminary investigations on miRNA:mRNA target pairs of interest have been generated. We used already available miRNA overexpression lines for analysis. Therefore, germination assays have been performed under various stress conditions including salt, mannitol, ABA, cold, heat or high light. Furthermore, putative miRNA targets have been quantified.

## **RESULTS**

The following section includes summaries, the both publications, that are part of this thesis and additional unpublished work.

### **Publication I: Impact of small RNAs in retrograde signalling pathways in *Arabidopsis thaliana***

Kristin Habermann<sup>1</sup>, Bhavika Tiwari<sup>1</sup>, Maria Krantz<sup>2</sup>, Stephan O. Adler<sup>2</sup>, Edda Klipp<sup>2</sup>, M. Asif Arif<sup>1\*</sup>, Wolfgang Frank<sup>1\*</sup>

- 1 Plant Molecular Cell Biology, Department Biology I, Ludwig-Maximilians-Universität München, LMU Biocenter, 82152 Planegg-Martinsried, Germany
- 2 Humboldt-Universität Berlin, Department Biologie, Bereich Theoretische Biophysik, 10115 Berlin, Germany


<https://doi.org/10.1111/tpj.14912>

#### **Abstract**

Chloroplast perturbations activate retrograde signalling pathways causing dynamic changes of gene expression. Besides transcriptional control of gene expression different classes of small non-coding RNAs (sRNAs) act in gene expression control, but comprehensive analyses regarding their role in retrograde signalling is lacking. We performed sRNA profiling in response to norflurazon (NF) that provokes retrograde signals in *A. thaliana* wild type and the two retrograde signalling mutants *gun1* and *gun5*. The RNA samples were also used for mRNA and long non-coding RNA (lncRNA) profiling to link altered sRNA levels to changes of their cognate target RNAs. We identified 122 sRNAs from all known sRNA classes that were responsive to NF in wild type. Strikingly, 140 and 213 sRNAs were found to be differentially regulated in both mutants indicating a retrograde control of these sRNAs. Concomitant with the changes in sRNA expression we detected about 1500 differentially expressed mRNAs in the NF treated wild type and around 900 and 1400 mRNAs that were differentially regulated in the *gun1* and *gun5* mutant with a high proportion (~30%) of genes

encoding plastid proteins. Furthermore, around 20% of predicted miRNA targets code for plastid localised proteins. The analyses of sRNA-target pairs identified pairs with an anticorrelated expression as well pairs showing other expressional relations pointing to a role of sRNAs in balancing transcriptional changes upon retrograde signals. Based on the comprehensive changes in sRNA expression we assume a considerable impact of sRNAs in retrograde-dependent transcriptional changes to adjust plastidic and nuclear gene expression.

# Identification of small non-coding RNAs responsive to *GUN1* and *GUN5* related retrograde signals in *Arabidopsis thaliana*

Kristin Habermann<sup>1</sup> , Bhavika Tiwari<sup>1</sup>, Maria Krantz<sup>2</sup>, Stephan O. Adler<sup>2</sup>, Edda Klipp<sup>2</sup>, M. Asif Arif<sup>1,\*</sup> and Wolfgang Frank<sup>1,\*</sup>

<sup>1</sup>Plant Molecular Cell Biology, Department Biology I, Ludwig-Maximilians-Universität München, LMU Biocenter, Planegg-Martinsried 82152, Germany, and

<sup>2</sup>Department Biologie, Bereich Theoretische Biophysik, Humboldt-Universität Berlin, Berlin 10115, Germany

Received 19 August 2019; revised 10 June 2020; accepted 17 June 2020; published online 8 July 2020.

\*For correspondence (e-mails wolfgang.frank@lmu.de; asif.arif@lmu.de).

## SUMMARY

Chloroplast perturbations activate retrograde signalling pathways, causing dynamic changes of gene expression. Besides transcriptional control of gene expression, different classes of small non-coding RNAs (sRNAs) act in gene expression control, but comprehensive analyses regarding their role in retrograde signalling are lacking. We performed sRNA profiling in response to norflurazon (NF), which provokes retrograde signals, in *Arabidopsis thaliana* wild type (WT) and the two retrograde signalling mutants *gun1* and *gun5*. The RNA samples were also used for mRNA and long non-coding RNA profiling to link altered sRNA levels to changes in the expression of their cognate target RNAs. We identified 122 sRNAs from all known sRNA classes that were responsive to NF in the WT. Strikingly, 142 and 213 sRNAs were found to be differentially regulated in both mutants, indicating a retrograde control of these sRNAs. Concomitant with the changes in sRNA expression, we detected about 1500 differentially expressed mRNAs in the NF-treated WT and around 900 and 1400 mRNAs that were differentially regulated in the *gun1* and *gun5* mutants, with a high proportion (~30%) of genes encoding plastid proteins. Furthermore, around 20% of predicted miRNA targets code for plastid-localised proteins. Among the sRNA–target pairs, we identified pairs with an anti-correlated expression as well pairs showing other expressional relations, pointing to a role of sRNAs in balancing transcriptional changes upon retrograde signals. Based on the comprehensive changes in sRNA expression, we assume a considerable impact of sRNAs in retrograde-dependent transcriptional changes to adjust plastidic and nuclear gene expression.

**Keywords:** small non-coding RNA, non-coding RNA, gene regulation, retrograde signalling, *gun1*, *gun5*, *Arabidopsis thaliana*.

## INTRODUCTION

Both mitochondria and chloroplasts are characteristic organelles of eukaryotes that have evolved through the endosymbiosis of distinct prokaryotic progenitors (Goksoyr, 1967). Cyanobacteria gave rise to plastids, and the majority of the endosymbiotic cyanobacterial genome was transferred into the nuclear DNA of the host organism. Consequently, most multiprotein complexes within the plastids are formed by organellar- and nuclear-encoded proteins, requiring a well-coordinated expression of both genomes (Zimorski *et al.*, 2014; Zhao *et al.*, 2019a). The nuclear gene expression is controlled by plastid-to-nucleus retrograde signalling (Kleine and Leister, 2016; Chan *et al.*, 2016), which is proposed to be mediated by several factors. For example, norflurazon (NF), a specific inhibitor of the

enzyme phytoene desaturase, which produces  $\beta$ -carotenoids from phytoene, causes repression of photosynthesis-associated nuclear genes (*PhANGs*) (Woodson *et al.*, 2011). Carotenoids are part of the light-harvesting complexes and protect the cells from photooxidative damage (Kim and Apel, 2013). In the presence of NF the chloroplast suffers from photooxidation, leading to characteristic bleaching symptoms of the green plant tissues caused by the degradation of chlorophyll (Breitenbach *et al.*, 2001). Several decades ago *Arabidopsis thaliana* mutant screens were performed to identify factors which specifically block the expression of *PhANGs* under conditions of chloroplast developmental prevention (Susek *et al.*, 1993; Mochizuki *et al.*, 2001; Meskauskiene *et al.*, 2001; Larkin *et al.*, 2003; Gray *et al.*, 2003; Gutierrez-Nava *et al.*, 2004; Ball *et al.*,

2004; Rossel *et al.*, 2006; Saini *et al.*, 2011). Several GENOME UNCOUPLED (*gun*) mutants were identified with disturbed retrograde signalling leading to a de-repression of *PhANGs*. Interestingly, five different *gun* mutants, *gun2* to *gun6*, are affected in the tetrapyrrole biosynthesis pathway (TPB). The *gun5* mutant has a defective regulatory CHLH subunit of the magnesium-chelatase (Mochizuki *et al.*, 2001). The *gun4* mutation also affects the subunit of the magnesium-chelatase, leading to an increased efficiency. The *gun2*, *gun3* and *gun6* mutants are impaired in heme oxygenase, phytylchromobilin synthase and Fe-chelatase, respectively (Woodson *et al.*, 2011; Woodson *et al.*, 2013). Based on these studies, it has been proposed that chloroplast metabolites may act as retrograde signals (Kakizaki *et al.*, 2009). The *gun1* mutant is not related to the remaining *gun* mutants since *GUN1* encodes a member of the chloroplast-localised pentatricopeptide repeat proteins, which usually act in post-transcriptional processes (Tadini *et al.*, 2016). The *gun1* mutant is able to perceive signals from the TPB, plastid gene expression and redox state, but the mode of action of *GUN1* in retrograde signalling remains unknown (Kleine and Leister, 2016). Microarray studies have been performed to compare transcriptional changes of *A. thaliana* wild type (WT) and *gun1* and *gun5* mutants in response to NF, revealing a strong correlation between the *gun1* and the *gun5* mutant because a large number of genes were consistently regulated in both mutants, including de-repression of *PhANGs*.

To date, all studies analysing gene expression in various retrograde signalling mutants focused on the analysis of protein-coding genes. However, it is well known that classes of non-coding RNAs (ncRNAs), including long ncRNAs (lncRNAs) as well as small ncRNAs (sRNAs), have important functions in diverse biological processes because they mainly act in the control of gene expression (Wang and Chekanova, 2017; Huang *et al.*, 2019).

lncRNAs with a size larger than 200 nucleotides were shown to have important functions in the control of gene expression (Wierzbicki *et al.*, 2008; Dinger *et al.*, 2009) and to exert their function by various mechanisms. One specific role of lncRNAs is the regulation of mRNA splicing, where they can either activate or inhibit specific splicing events (Ma *et al.*, 2014). They also mediate epigenetic modifications and act in microRNA (miRNA) target mimicry, where the lncRNA harbours a miRNA binding site, causing miRNA binding and sequestration (Franco-Zorrilla *et al.*, 2007; Swiezewski *et al.*, 2009; Heo and Sung, 2011). A specific gene regulatory class of ncRNA comprises sRNAs with a size of 20–24 nucleotides. They can interfere with nuclear transcription by regulating epigenetic modifications (Khraiweh *et al.*, 2010; Bannister and Kouzarides, 2011; Holoch and Moazed, 2015) or they can act post-transcriptionally by targeting RNAs, mediating RNA cleavage or translational inhibition (Meister and Tuschl, 2004; Bartel,

2004; Kim, 2005). sRNAs can be divided into two classes on the basis of their origin: hairpin RNA (hpRNA) and small interfering RNA (siRNA) (Axtell, 2013a). One of the most important classes of hpRNA are miRNAs, which are processed from stem-loop transcripts by DICER-LIKE1 enzymes (Park *et al.*, 2002; Meyers *et al.*, 2008) and guided through ARGONAUTE1 and RNA-induced silencing complex to their target RNAs by sequence complementarity to mediate their cleavage or translational inhibition (Wierzbicki *et al.*, 2008; Voinnet, 2009). Until now only one recent study reported on a functional role of miRNAs in retrograde signalling (Fang *et al.*, 2018). It was shown that tocopherols positively regulate the accumulation of 3'-phosphoadenosine 5'-phosphate (PAP), which is an inhibitor of exonuclease 2 (XRN2), which negatively regulates mRNA and pri-miRNA levels by degradation of 5' uncapped mRNA. Moreover, miR395 mediates cleavage of the mRNA encoding ATP sulfurylase (APS), the enzyme catalysing the initial step of PAP synthesis (Fang *et al.*, 2018).

Two other sRNA classes are formed from double stranded RNAs (dsRNAs), which are derived from endogenous transcripts and generate natural antisense transcript-derived siRNA (nat-siRNA) or trans-acting siRNA (ta-siRNA), based on their specific biogenesis pathways. Nat-siRNAs are generated from two genes encoding overlapping transcripts in antisense orientation, leading to the formation of dsRNA molecules (Borsani *et al.*, 2005). Nat-siRNAs are processed from these dsRNAs and mediate subsequent cleavage of one of the initial overlapping transcripts. According to their genomic location, NAT pairs can be distinguished into *cis*-NAT pairs, generated from opposing DNA strands within an identical genomic region, and *trans*-NAT pairs, produced from transcripts encoded by separated genomic regions (Lapidot and Pilpel, 2006; Yuan *et al.*, 2015). The first identified nat-siRNA was shown to have an important function in salt stress adaptation of *A. thaliana* (Borsani *et al.*, 2005), where it is involved in the regulation of proline biosynthesis. Unlike nat-siRNAs, ta-siRNA generation is triggered by miRNAs, since ta-siRNA precursor transcripts are cleaved in a miRNA-dependent manner and further processed into phased 21 nt ta-siRNA duplexes to control target RNAs (Chen, 2009). The role of sRNAs in retrograde signalling has not been analysed yet and information on the role of lncRNAs in retrograde control is completely lacking. To gain information whether these classes of ncRNA act in retrograde signalling, we made use of two well-characterised mutants affecting plastid-to-nucleus signalling events. *A. thaliana gun1* and *gun5* mutants were grown under standard conditions and in the presence of NF, and RNA expression profiles were compared to WT controls to identify functional sRNA–RNA target pairs that are modulated by retrograde signals.

## RESULTS

### De novo sRNA sequencing after norflurazon treatment

To identify sRNAs that may act in retrograde signalling pathways, seedlings of *A. thaliana* WT and the two retrograde signalling mutants *gun1* and *gun5* were treated for 4 days with 5  $\mu$ M NF under continuous light (Figure S1a) and sRNA sequencing was performed from six independent biological replicates samples, yielding a minimum of 5 million reads per replicate. The length distribution of all sRNA reads was analysed and we observed an enrichment of reads with a length of 21 and 24 nt (Figure S1b–d). The 21 nt peak corresponds to an expected enrichment of miRNAs, ta-siRNAs and nat-siRNAs, whereas the 24 nt peak complies with enriched repeat-associated sRNAs. The ShortStack sRNA analysis software has been used to map the sRNA data set against different reference databases (Table S1).

DeSeq2 was used to calculate the differential expression (2-fold regulation and false discovery rate [FDR]  $\leq 0.05$ ) of sRNAs between the samples, with a special focus on sRNAs that were differentially expressed in NF-treated samples with respect to their untreated controls, and in NF-treated *gun* mutants compared to the NF-treated WT (Table S2). Specific sRNA clusters arising from different ncRNA classes were found to be differentially expressed (Figure 1). These classes include mature miRNAs, *cis*-nat-siRNAs and *trans*-nat-siRNA, as well as sRNAs derived from lncRNAs. Upon growth on normal media, we identified only a small number of differentially regulated sRNAs in the *gun* mutants as compared to the WT, whereas the number of differentially regulated sRNAs between the mutants and WT strongly increased upon NF treatment (Table S3).

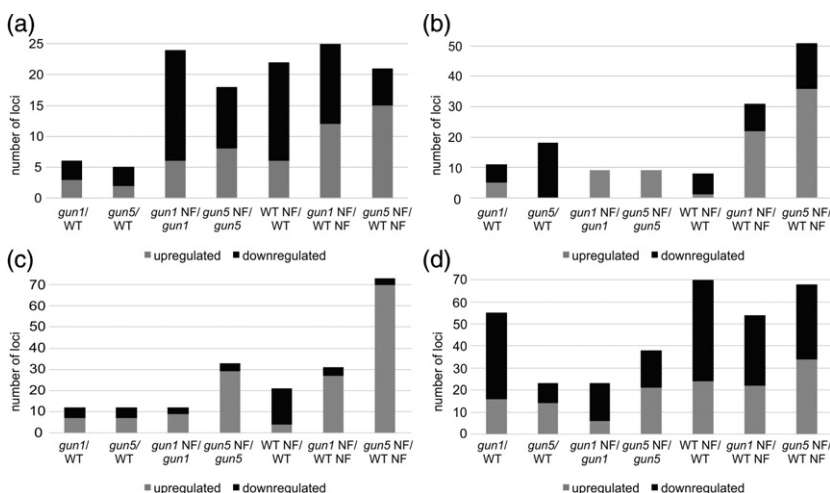
NF treatment caused an increased number of differentially expressed sRNAs in WT and both *gun* mutants, indicating a considerable sRNA regulation by retrograde signals. Furthermore, we observed a higher number of

differentially expressed sRNAs in both NF-treated *gun* mutants compared to NF-treated WT, pointing to a strong regulation of sRNAs that underlies specific retrograde signals in these mutants. Most of the changes affect miRNA and nat-siRNA expression levels, and we mainly focused on these sRNA classes with regard to their differential expression and further target analysis to predict the regulatory functions of these sRNAs (Figure 1 and Table S3).

### Analysis of differentially expressed miRNAs

Because beside a recent analysis of tocopherol-responsive miRNAs (Fang *et al.*, 2018) little is known about the role of miRNAs in retrograde signalling, we analysed changes in miRNA expression in response to NF in *A. thaliana* WT and in the *gun1* and *gun5* mutants. The comparison of differentially expressed miRNAs between the samples is shown in a hierarchically clustered heatmap (Figure 2a). Only a low number of differentially expressed miRNAs was observed in the untreated mutants compared to the WT control. In the *gun1* mutant (*gun1*/WT), only six differentially expressed miRNAs were detected, and only five miRNAs were detected in the untreated *gun5* mutant compared to the WT control (*gun5*/WT).

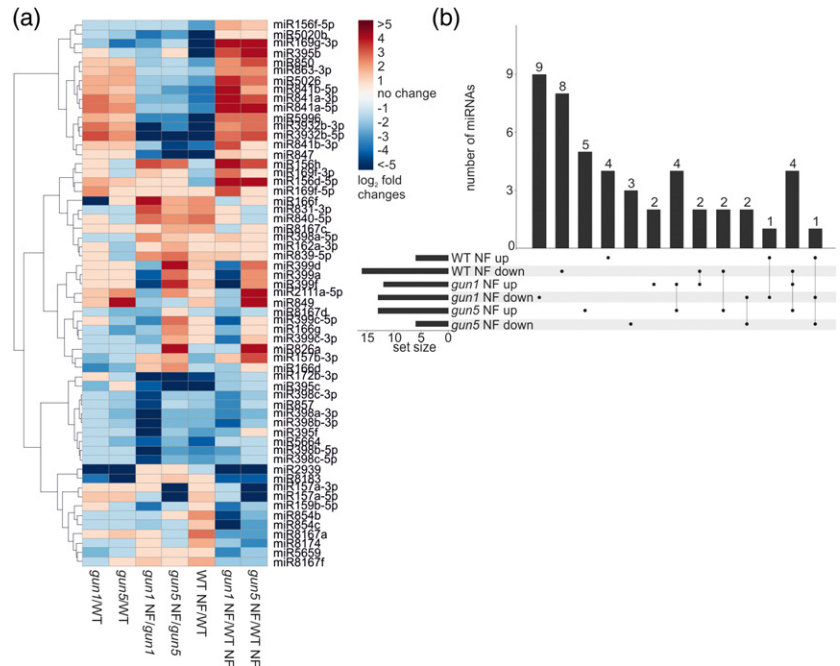
We hypothesised that miRNAs can play a role in retrograde signalling that should be reflected by an enrichment of differentially expressed miRNAs after NF treatment. Indeed, we observed a remarkable increase in the number of differentially expressed miRNAs in response to NF treatment with a similar number of NF-responsive miRNAs in the three analysed genotypes (Figure 2b). In total, we observed 22 miRNAs to be differentially regulated in the NF-treated WT compared to the untreated control (WT NF/WT). Twenty-four miRNAs were differentially expressed in the NF-treated *gun1* compared to the untreated *gun1* mutant (*gun1* NF/*gun1*), and 18 miRNAs were differentially regulated in the NF-treated *gun5* mutant compared to the untreated *gun5* control (*gun5* NF/*gun5*).



**Figure 1.** Differentially expressed sRNAs within the different samples. Overview of differentially regulated sRNAs between the different samples ( $\log_2(\text{FC}) \leq -2$  or  $\geq +2$ ;  $\text{FDR} \leq 0.05$ ) subdivided into specific sRNA classes. (a) miRNAs, (b) sRNAs derived from lncRNA, (c) *cis*-NAT pairs and (d) *trans*-NAT pairs. The up- and downregulation of the members of each class are depicted by grey (up) and black (down) partitions of the respective bars.



**Figure 2.** Behaviour of the differentially expressed miRNAs. (a) Hierarchically clustered (UPGMA) heat-map depicting miRNAs that are differentially regulated in at least one sample displaying normalised  $\log_2(\text{FC})$  values. (b) UpSet plot depicting the number of differentially expressed miRNAs in response to NF in WT (WT NF/WT) and both *gun* mutants (*gun1* NF/WT NF and *gun5* NF/WT NF).



Interestingly, we further detected miRNAs that seem to be controlled by retrograde signals, as de-repressed miRNAs were observed in the *gun1* and *gun5* mutants in response to NF treatment, which is reminiscent of the de-repression of *PhANGs* in these mutants. We focused on miRNAs with altered expression levels in the treated WT (WT NF/WT) and correlated them with differentially expressed miRNAs in the NF-treated *gun* mutants (Figure 2b). In response to NF, two miRNAs (miR169g-3p and miR5996) showed patterns of de-repression in both *gun* mutants similar to de-repressed *PhANGs*, and one miRNA (miR3932-5p) was de-repressed only in the NF-treated *gun5* mutant compared to the treated WT (*gun5* NF/WT NF). Furthermore, five miRNAs were downregulated in the treated WT (WT NF/WT) and were upregulated in at least one NF-treated *gun* mutant (*gun* NF/WT NF). We also identified two miRNAs which seemed to be controlled by retrograde signals in an opposite manner. These two miRNAs were found to be upregulated in the treated WT (WT NF/WT) and downregulated in at least one of the treated *gun* mutants (*gun* NF/WT NF). In addition, we found miRNAs which showed a specific regulation restricted to NF-treated *gun* mutants when compared to the treated WT. Two miRNAs were found to be downregulated in both treated *gun* mutants (*gun* NF/WT NF). Moreover, nine miRNAs were specifically downregulated in the NF-treated *gun1* mutant compared to the treated WT (*gun1* NF/WT NF), and the expression of three miRNAs was reduced in the treated *gun5* mutant (*gun5* NF/WT NF). Furthermore, two miRNAs were upregulated in the treated *gun1* mutant and five miRNAs were upregulated in the treated *gun5* mutant (*gun* NF/

WT NF). We also detected four upregulated miRNAs common for both treated *gun* mutants (*gun* NF/WT NF).

#### Differentially regulated nat-siRNAs

To identify nat-siRNAs from our sRNA sequencing data we made use of different accessible databases (Table S1) comprising experimentally validated and computationally predicted *cis*- and *trans*-NAT pairs (Jin *et al.*, 2008; Zhang *et al.*, 2012; Yuan *et al.*, 2015).

We identified 12 various *cis*-NAT pairs producing differentially regulated nat-siRNA clusters in both untreated *gun1* and *gun5* mutants (*gun1*/WT and *gun5*/WT) (Figure 1c). Besides this, 57 and 23 *trans*-NAT pairs were detected to produce differentially regulated nat-siRNA in the *gun1* and *gun5* mutants, respectively (Figure 1d).

Upon NF treatment we detected 21 *cis*-NAT pairs (Figure S2a) and 70 *trans*-NAT pairs (Figure S2b) in the WT (WT NF/WT) producing differentially regulated nat-siRNA clusters from at least one transcript of these NAT pairs. In the treated *gun1* mutant (*gun1* NF/*gun1*), nat-siRNAs from 12 *cis*-NAT pairs were detected to be differentially expressed (Figure 1c) and 23 differentially regulated *trans*-NAT pairs producing nat-siRNA clusters were identified to be differentially regulated in the treated *gun1* mutant (*gun1* NF/*gun1*). In the NF-treated *gun5* mutant, we identified 33 *cis*-NATs and 38 *trans*-NATs generating differentially expressed nat-siRNAs (*gun5* NF/*gun5*) (Figure 1c,d).

The overlap and co-regulation as well as specific expression of the differentially expressed *cis*-NAT pairs and *trans*-NAT pairs producing differentially regulated nat-siRNA clusters were analysed between the samples and



are shown in an UpSet plot (Figure S2). We focused on the analysis of NF-responsive differentially expressed nat-siRNAs in the WT to provide information on nat-siRNAs that are controlled by retrograde signals. Moreover, we compared NF-treated WT with both NF-treated *gun* mutants to identify NF-responsive nat-siRNA misregulation that is caused by the perturbed retrograde signals in these mutants. We detected 31 *cis*-NAT pairs and 54 *trans*-NAT pairs producing differentially expressed nat-siRNAs in the NF-treated *gun1* mutant (*gun1* NF/WT NF). In the NF-treated *gun5* mutant we detected 73 *cis*-NAT pairs and 68 *trans*-NAT pairs that produce differentially regulated nat-siRNA clusters (*gun5* NF/WT NF). For both *gun* mutants the majority of *cis*-derived nat-siRNAs were upregulated, whereas the majority of *trans*-derived nat-siRNAs were downregulated in these mutants (Figure S2). We identified five *cis*-NAT pairs to be downregulated in the treated WT (WT NF/WT) and upregulated in both treated *gun* mutants (*gun* NF/WT NF), thus representing the *gun*-specific de-repression of nuclear-encoded *PhANGs* (Figure S2a). We also detected one nat-siRNA cluster produced from a *cis*-NAT pair displaying an opposing expression pattern (upregulated in WT NF/WT and downregulated in both *gun* NF/WT NF). Within *trans*-derived nat-siRNAs we identified 19 sRNA clusters that were differentially regulated in response to NF in WT (WT NF/WT) and showed further differential regulation in response to NF in both *gun* mutants (*gun* NF/WT NF) (Figure S2b). Five of them resemble the *gun*-specific de-repression, since they were downregulated in the WT (WT NF/WT) and upregulated in both mutants (*gun* NF/WT NF). Eleven *trans*-derived nat-siRNAs displayed an opposite expression and were upregulated in the treated WT (WT NF/WT) and downregulated in both treated *gun* mutants (*gun* NF/WT NF). In addition, two differentially expressed nat-siRNA from *trans*-NAT pairs were downregulated within all three samples and another one was upregulated in the treated WT (WT NF/WT) and in the treated *gun5* mutant (*gun5* NF/WT NF) and downregulated in the treated *gun1* mutant (*gun1* NF/WT NF).

#### Other differentially regulated sRNA classes

Besides the differentially expressed miRNAs and NAT pairs we also found differentially expressed sRNAs produced from lncRNAs and phased siRNA (phasiRNAs) precursors (Figure S3 and Table S3). Similar to miRNAs and nat-siRNAs, we detected only a small number of differentially regulated sRNA clusters derived from lncRNA precursors in the untreated genotypes (*gun*/WT). In total, 11 differentially expressed sRNA clusters produced from lncRNAs were noticed in the *gun1* mutant (*gun1*/WT) and all 18 differentially expressed sRNA clusters in the untreated *gun5* mutant were downregulated (*gun5*/WT).

When comparing the individual genotypes with and without NF treatment, we identified only a considerably

small number of differentially expressed sRNAs. In the treated WT, eight sRNA clusters processed from lncRNA precursors were identified to be differentially expressed (WT NF/WT). For both treated *gun* mutants, we observed nine different upregulated sRNA clusters (*gun* NF/*gun*).

Comparing the NF-treated *gun* mutants with the NF-treated WT we noticed an increase in the number of differentially expressed sRNAs (Figure S3). Generally, we observed a higher number of upregulated sRNA clusters produced from lncRNA precursors in both treated *gun* mutants (*gun* NF/WT NF). Of 31 differentially expressed sRNA clusters, 22 were detected to be upregulated in the NF-treated *gun1* mutant (*gun1* NF/WT NF). For the treated *gun5* mutant, 36 out of 51 differentially expressed sRNA clusters were found to be upregulated (*gun5* NF/WT NF). We detected only one sRNA cluster derived from a lncRNA precursor that was downregulated in the treated WT (WT NF/WT) and upregulated in both NF-treated *gun* mutants (*gun* NF/WT NF). Two sRNA clusters were downregulated in the treated WT (WT NF/WT) and upregulated in the treated *gun1* mutant (*gun1* NF/WT NF). Another two sRNA clusters derived from lncRNA precursors were downregulated in the treated WT (WT NF/WT) and upregulated in the treated *gun5* mutant (*gun5* NF/WT NF). Furthermore, 11 sRNA clusters produced from lncRNA precursors were similarly regulated in both treated *gun* mutants (*gun* NF/WT NF), with six of them upregulated and five downregulated (Figure S3).

In addition to the lncRNA-derived sRNA clusters, we identified two differentially expressed phasiRNAs. One phasiRNA derived from locus AT1G63070 was 5.8-fold upregulated in the untreated *gun1* mutant (*gun1*/WT) and 4.1-fold upregulated in the NF-treated *gun1* mutant (*gun1* NF/WT NF). The second phasiRNA produced from the locus AT5G38850 was 2.6-fold downregulated in the treated WT (WT NF/WT) and 3.5-fold downregulated in the treated *gun1* mutant (*gun1* NF/*gun1*).

#### Analysis of lncRNA and mRNA in the *gun* mutants

Besides sRNA sequencing, we also sequenced mRNAs and lncRNAs to gain more information about NF-dependent regulation of lncRNAs and to examine the correlation of sRNAs with their targets.

The samples were mapped against the *A. thaliana* genome deposited in Araport11 (Table S4) and differential expression of mRNA and lncRNA between the samples (Tables S5 and S6) was calculated with Cuffdiff. Representative transcripts belonging to different RNA classes showing differential expression levels in the RNA sequencing data were selected for expression analyses by quantitative real-time PCR (qRT-PCR), which confirmed the mRNA and lncRNA sequencing data (Figure S4). We selected various genes which were detected to be differentially expressed in the treated WT as well as in both NF-treated *gun* mutants. Furthermore, we selected two transcripts each

that displayed a low, moderate and high abundance, respectively. In addition, we included one lncRNA that was found to be differentially regulated in all three samples.

#### Classification of differentially expressed ncRNAs detected via ribosomal depleted RNA sequencing

We identified differentially expressed transcripts belonging to distinct ncRNA classes (Tables S5 and S6), including lncRNAs, which may act in regulatory processes of gene expression, as well as tRNA, rRNA and small nucleolar RNA (snoRNA), which act in protein translation and splicing and usually have few regulatory functions. Under normal growth conditions we identified 10 differentially expressed ncRNAs in each of the *gun* mutants as compared to the untreated WT (Table 1 and Figure S5a). The number of differentially expressed ncRNAs increased upon NF treatment, indicating potential roles upon plastid perturbations that trigger retrograde signalling. In total, we identified 34 differentially expressed ncRNAs in the NF-treated WT compared to the untreated control (Table 1). In the NF-treated *gun1* and *gun5* mutants (*gun* NF/*gun*), we identified 32 and 70 differentially expressed ncRNAs, respectively. Interestingly, in the NF-treated *gun* mutants we observed 20 and 45 differentially expressed ncRNAs in the *gun1* and *gun5* mutants (*gun* NF/WT NF), respectively.

An UpSet plot (Figure S5b) depicts the distribution of differentially expressed ncRNAs between various samples (WT NF/WT, *gun1* NF/WT NF and *gun5* NF/WT NF). We identified two interesting lncRNAs (AT1G05562 and AT4G13495), which represent the classical *gun*-related expression as these show a downregulation in response to NF treatment in WT, but are upregulated in both NF-treated *gun* mutants. Furthermore, three lncRNAs (AT3G01835, AT5G07325 and AT5G07745) were identified to be upregulated in the NF-treated WT (WT NF/WT) and downregulated in the treated *gun1* mutant (*gun1* NF/WT NF).

Another interesting lncRNA (AT4G13495) was de-repressed in both NF-treated *gun* mutants with 7.2-fold and

3.4-fold upregulation in *gun1* and *gun5* mutants (*gun* NF/WT NF), respectively, whereas this lncRNA was highly downregulated (fold change [FC] of  $-10.5$ ) in the treated WT (WT NF/WT). From our sRNA data we already detected sRNAs arising from this lncRNA and in agreement with the expression level of this lncRNA, the total sRNAs generated from this transcript were downregulated in the treated WT (FC of  $-2.8$ ; WT NF/WT) and 3.4-fold upregulated in the treated *gun1* mutant (*gun1* NF/WT NF). Interestingly, this lncRNA overlaps with three individual miRNA precursors (miR5026, miR850 and miR863) in sense direction, suggesting that these miRNAs can be processed from the individual precursors as well as from the overlapping lncRNA. In line with this hypothesis, we observed a consistent differential expression of the lncRNA and the three individual miRNAs within the analysed samples (Table 2).

In addition, we identified two differentially regulated lncRNAs overlapping with mRNA transcripts in antisense that may act as precursors for the generation of nat-siRNAs. One lncRNA (AT1G05562) that may act as a natural antisense transcript was downregulated in the treated WT (FC of  $-3.7$ ; WT NF/WT) and upregulated in the treated *gun1* and *gun5* mutants with a FC of 3.9 and 4.2 (*gun* NF/WT NF), respectively. This lncRNA transcript is able to overlap with an mRNA encoding an UDP-glucose transferase (AT1G05560). Furthermore, the overlapping mRNA transcript was downregulated in the treated WT (FC of  $-5.8$ ; WT NF/WT) and upregulated in both treated mutants (FC of 3.6 for *gun1* NF/WT NF; FC of 3.1 for *gun5* NF/WT NF). We also detected differentially expressed sRNA clusters processed from this region in the sRNA sequencing data in the treated WT (FC of  $-4.4$  for WT NF/WT) as well as in the treated *gun1* mutant (FC of 6.6 for *gun1* NF/WT NF). Thus, the regulation of nat-siRNAs correlates with the expression of the respective lncRNA-mRNA transcript pair and the differential expression seems to be regulated by specific retrograde signalling pathways.

**Table 1** Overview of differentially expressed ncRNAs in response to NF in *A. thaliana* WT and *gun1* and *gun5* mutants

	<i>gun1</i> / WT	<i>gun5</i> / WT	WT NF/ WT	<i>gun1</i> NF/ <i>gun1</i>	<i>gun5</i> NF/ <i>gun5</i>	<i>gun1</i> NF/ WT NF	<i>gun5</i> NF/ WT NF
lncRNAs	6	5	15	13	34	11	20
snRNAs	0	0	3	0	8	1	5
snoRNAs	0	0	3	3	11	2	4
rRNAs	0	0	2	0	1	1	0
tRNAs	0	0	3	4	1	2	1
pseudogenes	4	5	5	6	9	2	10
transcript regions	0	0	2	5	3	1	4
MIR precursors	0	0	1	1	2	0	0
antisense RNAs	0	0	0	0	1	0	1
Total	10	10	34	32	70	20	45

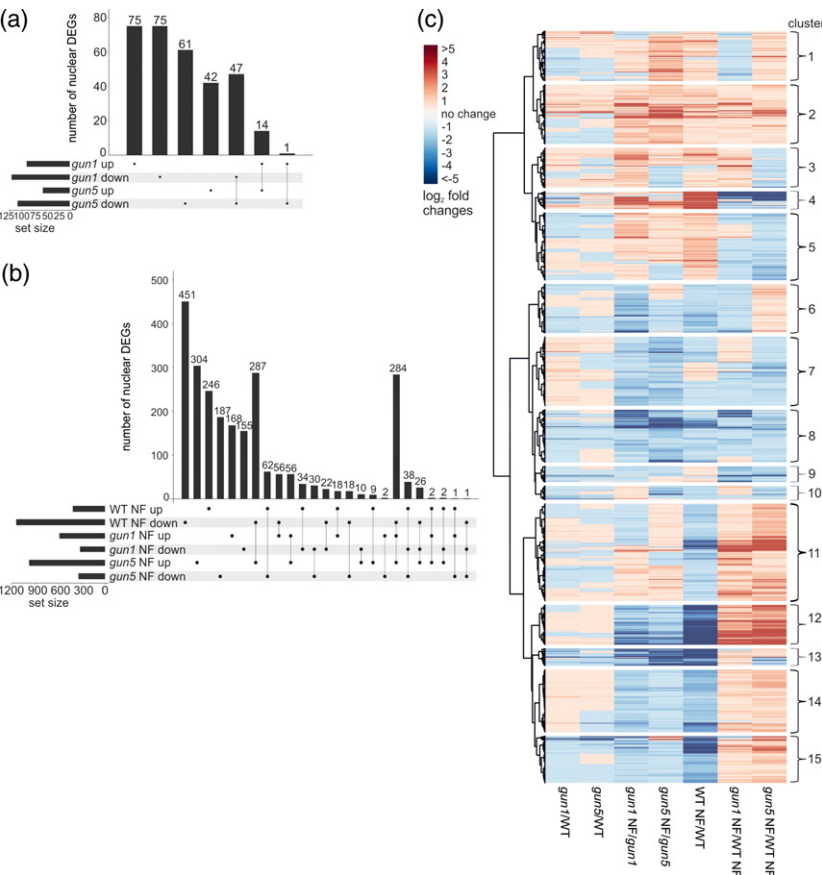
**Table 2** Expression data for the lncRNA AT4G13495 that overlaps in sense with the individual miRNA precursors miR5026, miR850 and miR863

ID	FC WT NF/WT	FDR	FC <i>gun1</i> NF/WT		FC <i>gun5</i> NF/WT	
			NF	FDR	NF	FDR
AT4G13495	−10.54	0.001	7.17	0.001	3.36	0.001
miR5026	−2.63	0.085	4.08	0.004	3.08	0.042
miR850	−2.58	0.077	2.61	0.068	3.81	0.007
miR863-5p	−1.7	0.561	1.29	0.941	2.34	0.467
miR863-3p	−1.51	0.538	2.7	0.025	2.61	0.045

In addition, we identified a *TAS3* precursor transcript (AT3G17185) that was downregulated in the NF-treated WT (FC of −2.9 for WT NF/WT) and de-repressed in the treated *gun5* mutant (FC of 2.6 for *gun5* NF/WT NF). Ta-siRNAs produced from the *TAS3* transcript control the expression of transcripts coding for auxin response factors such as *ARF2*, *ARF4* and *ETT*. However, we detected neither differentially expressed *TAS3*-derived ta-siRNAs nor differential expression of their cognate targets between the analysed samples.

# Differentially regulated nuclear- and organellar-encoded mRNAs after NF treatment

In parallel to sRNA and lncRNA, we analysed the data obtained from the ribosomal depleted nuclear- (Figure 3) and organellar-encoded (Figure 4) RNA sequencing to identify protein-coding mRNAs that are regulated by retrograde signalling pathways. Furthermore, to categorise putative functions of differentially regulated RNAs after NF treatment, Gene Ontology (GO) enrichment terms were explored (Table S8 and Figure S6). We detected only a low number of differentially expressed genes (DEGs), with 212 and 165 differentially expressed transcripts in the untreated *gun1* and *gun5* mutants (*gun*/WT), respectively (Figure 3a). However, when we analysed differential gene expression in response to NF, we observed a remarkable increase in the number of DEGs (Figure 3b). We identified 1557 DEGs in the WT in response to NF (WT NF/WT). For both treated mutants compared to their respective untreated controls, we identified slightly lower numbers of DEGs. In total, 1361 DEGs were identified in the treated *gun1* mutant (*gun1* NF/*gun1*) and 1177 DEGs were detected in the treated *gun5* mutant (*gun5* NF/*gun5*). In addition, we compared mRNA expression between the NF-treated *gun* mutants and the NF-treated WT. We identified 905 DEGs in



**Figure 3.** Distribution of nuclear DEGs in the untreated and NF-treated samples. (a) UpSet plot showing the distribution of differentially regulated mRNAs in the untreated *gun* mutants compared to the WT. (b) UpSet plot depicting the distribution of differentially regulated mRNAs in response to NF in WT (WT NF/WT) and both *gun* mutants (*gun1* NF/WT NF and *gun5* NF/WT NF). (c) Hierarchically clustered (UPGMA) heatmap of normalised  $\log_2(FC)$  values from nuclear-encoded DEGs with 15 clusters based on co-expression patterns.

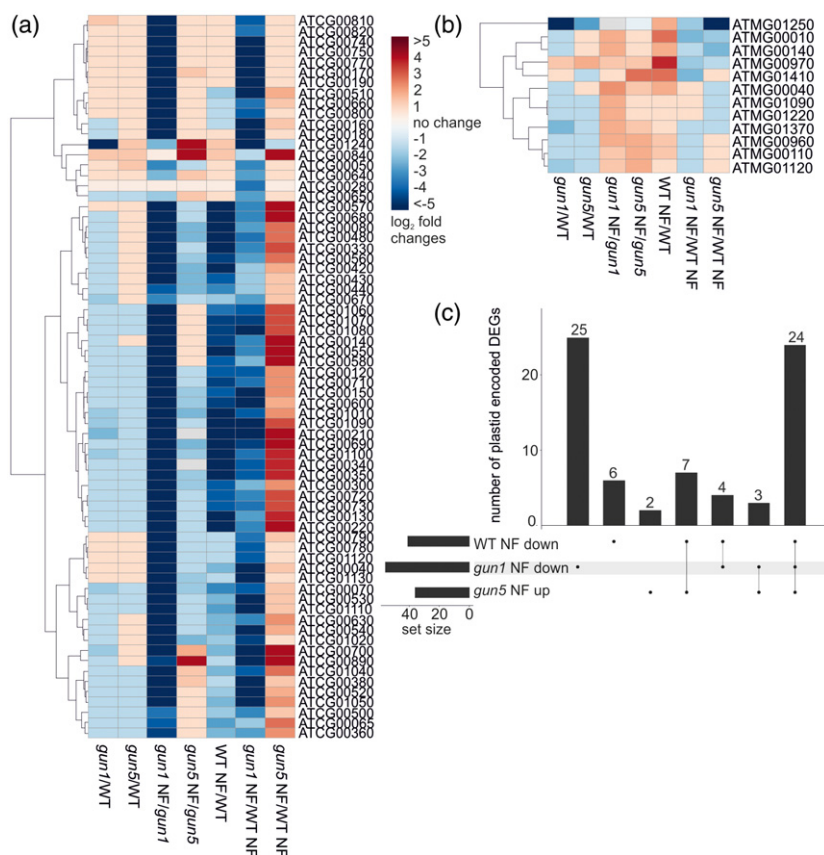
the NF-treated *gun1* mutant (*gun1* NF/WT NF) and 1319 DEGs in the treated *gun5* mutant (*gun5* NF/WT). We generated a hierarchically clustered heatmap from all 3352 nuclear-encoded mRNAs that were differentially regulated in at least one sample (Figure 3c). Based on the co-expression of DEGs we were able to separate 15 specific clusters of differentially regulated nuclear-encoded genes (Table S7). We identified 1557 DEGs in the treated WT (WT NF/WT) and 75% of the mRNAs were downregulated. As expected, the NF-treated *gun* mutants behaved in an opposite manner, as the majority of the RNAs were upregulated, with 65% and 75% upregulated DEGs in the treated *gun1* and *gun5* mutants (*gun* NF/WT NF), respectively.

To identify the most interesting candidates regulated by retrograde signals, we analysed the overlap between the treated WT (WT NF/WT) and both treated *gun* mutants (*gun* NF/WT NF) to detect those genes that display a typical *gun*-related expression in both mutants (Figure 3b). We identified 284 DEGs in response to NF in WT (WT NF/WT) as well as in both *gun* mutants (*gun* NF/WT NF). These DEGs seem to be controlled by retrograde signalling pathways, because they are repressed by NF in the WT and de-repressed in the *gun* mutants. Furthermore, we detected 56 DEGs with a specific de-repression in the treated *gun1* mutant (*gun1* NF/WT NF) and another 287 DEGs

specifically de-repressed in the *gun5* mutant (*gun5* NF/WT NF). Most likely, the regulation of the genes requires specific retrograde signals, as we identified genes showing a specific de-repression restricted to only one of the *gun* mutants.

Besides the analysis of nuclear-encoded genes, we investigated organellar gene expression and studied the expression of genes encoded by the plastidic and mitochondrial genomes in the WT and both *gun* mutants in the absence or presence of NF. We generated two hierarchically clustered heatmaps for plastidic (Figure 4a) and mitochondrial (Figure 4b) genes that were differentially expressed in at least one of the samples. As expected, we only detected eight mitochondrial genes with differential expression in at least one of the samples, as NF treatment affects carotenoid biosynthesis in the plastids and should not directly affect mitochondrial gene expression. Furthermore, the low number of affected genes in the mitochondria indicates an insignificant crosstalk between plastids and mitochondria triggered by plastid-derived retrograde signals.

In contrast, we detected a considerable high number of differentially regulated plastid-encoded genes. Upon growth in the absence of NF, none of the plastid-encoded genes were differentially expressed in the *gun5* mutant



**Figure 4.** Distribution of differentially expressed plastidic and mitochondrial DEGs in the untreated and NF-treated samples. Hierarchically clustered (UPGMA) heatmap depicting (a) plastidic and (b) mitochondrial genes that are differentially expressed in at least one of the samples displaying normalised log<sub>2</sub>(FC) values. (c) UpSet plot depicting the expression of plastid-encoded DEGs detected in the NF-treated *gun* mutants (*gun1* NF/WT NF and *gun5* NF/WT NF) and in the NF-treated WT (WT NF/WT).

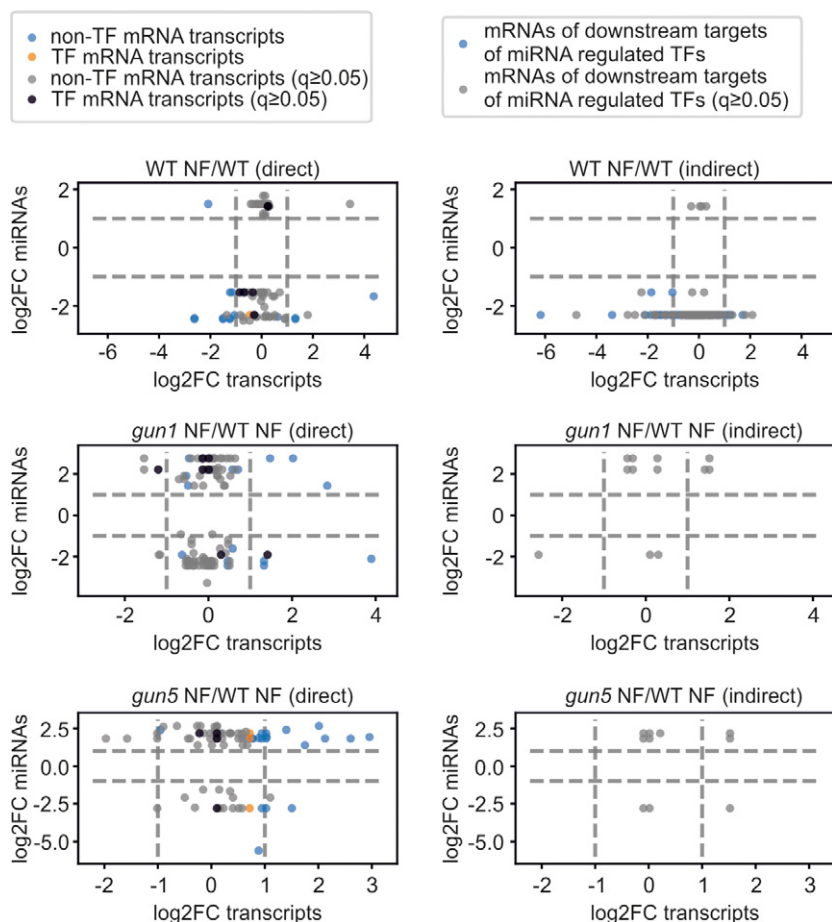


and only one plastid-encoded gene was differentially expressed in the *gun1* mutant (*gun*/WT). However, after NF treatment we detected 41, 56 and 36 differentially expressed plastid-encoded genes in the treated WT (WT NF/WT) and *gun1* and *gun5* mutants (*gun* NF/WT NF), respectively. Furthermore, we noticed a highly interesting phenomenon: Almost all plastid-encoded differentially expressed mRNAs were downregulated in the NF-treated *gun1* mutant (*gun1* NF/WT NF) and upregulated in the NF-treated *gun5* mutant (*gun5* NF/WT NF). Thus, based on the plastidic gene expression, both mutants respond in an almost completely opposed manner to NF treatment, suggesting specific perturbations in the NF-triggered organellar signalling pathways. We observed 27 differentially expressed plastid-encoded mRNAs in response to NF in both the *gun1* and the *gun5* mutant (Figure 4c) compared to the NF-treated WT, but they were regulated in an opposing manner: They were all downregulated in the treated *gun1* mutant, but upregulated in the treated *gun5* mutant.

### miRNA target analysis

We performed miRNA target prediction with 'psRNATarget' using all protein-coding and non-coding transcripts

from Araport11 to correlate the expression of miRNAs with putative target RNA transcripts (Table S9). For each predicted miRNA target, we considered its expression changes to subclassify the miRNA–RNA pairs. For the differentially regulated miRNAs, which were detected in WT NF/WT, *gun1* NF/WT NF and *gun5* NF/WT NF, we were able to predict 218 protein-coding targets as well as 16 non-coding target RNAs, and some of these can be targeted by several miRNAs. We generated a non-redundant list of miRNA targets and excluded transcripts with low fragments per kilobase of transcript per million reads (FPKM) values ( $\geq 5$ ). Applying these parameters, we obtained 119 predicted miRNA targets that were categorised into three different classes based on their expression. It has to be noted that a specific miRNA–RNA pair can be grouped into different categories since the miRNA as well as the cognate RNA target can be differently regulated between the analysed samples. The first category comprises miRNA–RNA pairs that are 'unchanged' according to the FC of the RNA transcript (but not the miRNA) and includes 101 miRNA–RNA pairs. The second category contains seven miRNA–RNA pairs that show an anticorrelated expression pattern, and the third category encompasses 16 miRNA–RNA pairs



**Figure 5.** Scatter plots of differentially expressed miRNAs and their targets. Only miRNAs with  $FDR \leq 0.05$  were included. The direct plots (left panel) depict all differentially expressed miRNAs and their direct predicted target transcripts. MiRNA target transcripts encoding transcription factors are shown in orange ( $FDR \leq 0.05$ ) and black ( $FDR \geq 0.05$ ). MiRNA target transcripts encoding other proteins are shown in blue ( $FDR \leq 0.05$ ) and grey ( $FDR \geq 0.05$ ). The indirect plots (right panel) depict downstream targets of transcription factors that are miRNA-regulated. Here, the mRNAs of these downstream genes are plotted against the miRNAs controlling their respective transcription factor mRNAs. The blue dots correspond to  $FDR \leq 0.05$  and the grey dots to  $FDR \geq 0.05$ .

where the miRNA and the predicted target show the same direction of their differential expression (both up- or both downregulated). Two different miRNAs together with at least two of their targets were validated by qRT-PCR, confirming their anticorrelated expression pattern (Figure S7a, b). Scatter plots (Figure 5) were created to show the distribution of the differentially regulated miRNAs and their correlating targets and were divided into 'direct' and 'indirect' scatter plots. The direct plots show the correlation of miRNAs and their cognate RNA targets either coding for transcription factors or coding for other proteins. The indirect scatter plots depict the expression of downstream genes that are controlled by miRNA-regulated transcription factors. From the direct scatter plot, it is obvious that most differentially expressed miRNA targets do not encode transcription factors. Nevertheless, we identified transcription factor transcripts which are controlled by miRNAs, and their effect on the transcription factor targets can be seen in the indirect plots. For example, the indirect plot shows many differentially expressed transcripts coding for transcription factors, which are controlled by miRNAs in the NF-treated WT (WT NF/WT).

We identified one miRNA–RNA target pair (Table S9) that has been shown to play a role in the acclimation to phosphate deficiency. MiR399a was downregulated in the treated *gun1* mutant (*gun1* NF/WT NF), whereas the expression of its target *PHO2* (AT2G33770), encoding a ubiquitin-conjugating E2 enzyme, remained unchanged in the treated *gun1* mutant (*gun1* NF/WT NF). MiR850 and its cognate target, encoding a chloroplast RNA-binding protein (AT1G09340), belong to the category of miRNA–target pairs showing the same expression (Table S9) since both were upregulated in the treated *gun5* mutant (*gun5* NF/WT NF). This chloroplast RNA-binding protein is necessary for the proper function of the chloroplast and mutations in this gene cause growth deficiency (Fettke *et al.*, 2011). Furthermore, we also identified miR157a-5p (FC of –7 in *gun5* NF/WT NF), displaying an anticorrelated expression to its target PHOTOSYSTEM II REACTION CENTRE PSB28 PROTEIN (AT4G28660), which is 2.9-fold upregulated (*gun5* NF/WT NF). PSB28 is highly conserved in photosynthetic eukaryotes and lack of *PSB28* results in a pale-green phenotype in rice, pointing to a role in the assembly of chlorophyll-containing proteins such as CP47 (Lu, 2016).

### Nat-siRNA target analysis

We detected a larger number of differentially expressed sRNAs arising from predicted NAT pairs than from any other sRNA class in the treated WT (WT NF/WT) as well as in both NF-treated *gun* mutants (*gun* NF/WT NF). Filtering the differentially expressed nat-siRNAs with at least five normalised reads in one of six samples (WT, WT NF, *gun1*, *gun1* NF, *gun5* and *gun5* NF) led to a total number of 73 non-redundant *cis*-NATs and 193 non-redundant *trans*-NAT

pairs. These pairs were further analysed and we only selected the nat-siRNA producing transcript pairs with at least five normalised reads for one of the two overlapping transcripts. This reduced the number to 64 non-redundant *cis*-NAT and 40 non-redundant *trans*-NAT pairs (Table S10). The expression changes of two nat-siRNAs together with their overlapping transcripts in NF-treated WT and the NF-treated *gun5* mutant were confirmed by qRT-PCR (Figure S7c,d).

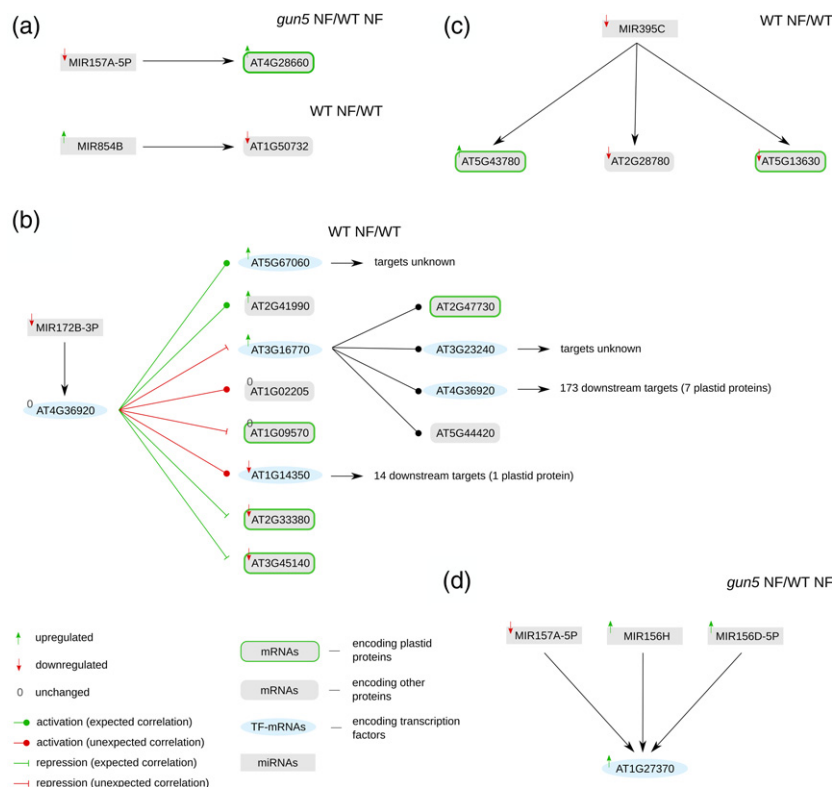
For many *trans*-NAT pairs, we observed that one of the transcripts was derived from a transposable element or a pre-tRNA, whereas the second overlapping transcript represented a protein-coding gene. Among these *trans*-NAT pairs, we only detected one overlapping transcript encoding a plastid-localised protein, suggesting a low impact of *trans*-NAT pairs in the adjustment of plastid and nuclear gene expression in response to NF. The *trans*-nat-siRNA generated from this pair was found to be downregulated in the treated WT (WT NF/WT) and upregulated in both treated *gun* mutants (*gun* NF/WT NF). The first overlapping transcript codes for the plastid-localised UDP-glucosyl transferase 75B2 (AT1G05530), which is able to bind UDP-glucose, important for cellulose and callose synthesis (Hong *et al.*, 2001). Its expression was unchanged in the treated WT (WT NF/WT) as well as in both treated *gun* mutants (*gun* NF/WT NF). The second overlapping transcript represented a lncRNA (AT1G05562) that was downregulated in the treated WT (WT NF/WT) and upregulated in both treated *gun* mutants (*gun* NF/WT NF).

Interestingly, out of 64 *cis*-NAT pairs that give rise to differentially regulated nat-siRNAs, we detected 31 individual transcripts which encode plastid proteins, indicating a considerable role of *cis*-NAT pairs in the direct control of genes coding for plastid proteins via NF-triggered retrograde signals. Moreover, within the *cis*-NAT pairs we identified 35 individual transcripts encoding nuclear-localised proteins, pointing to a large impact of these in the indirect adjustment of nuclear gene expression via nuclear regulatory proteins. One sRNA processed from a *cis*-NAT pair was detected to be downregulated in the treated WT (WT NF/WT). Interestingly, both overlapping transcripts were identified to encode plastid-localised proteins. The expression of the first overlapping transcript (AT1G29900), which codes for a subunit of carbamoyl phosphate synthetase, which is presumed to be necessary for the conversion of ornithine to citrulline in the arginine biosynthesis pathway (Molla-Morales *et al.*, 2011), was unchanged (WT NF/WT). In agreement with the expression of the nat-siRNAs, the second overlapping transcript (AT1G29910) was downregulated by NF in the WT (WT NF/WT). This transcript encodes a chlorophyll A/B-binding protein, which is the major protein of the light-harvesting complex and is required for absorbing light during photosynthesis.

## Network analysis

In order to gain a comprehensive picture of the role of miRNAs in retrograde signalling and to analyse possible downstream effects, we investigated a miRNA–RNA–target network that also comprises related transcription factor to target gene connections. The results were combined in a complex interaction network (Data S1), since one miRNA can control many mRNAs encoding transcription factors, which in turn control several downstream genes, but also one miRNA target can be controlled by numerous miRNAs (Figure S8 and Table S12). Within the considered network, most miRNAs regulate just a small number of target transcripts (Figure S8a), but there are some miRNAs regulating up to 140 targets. In contrast, the majority of miRNA targets are regulated by only a few miRNAs, but there are still some targets that can be regulated by up to 15 miRNAs (Figure S8b). We observed that miRNAs controlling the highest number of targets mainly regulate mRNAs that do not encode transcription factors (Figure S8c), while the distribution of miRNA targets encoding transcription factors indicates that most miRNAs regulate only a small number of such targets, with the highest number being eight (Figure S8d). Some motifs are recurrent in the miRNA–RNA target network (Figure 6). We explored the network for different characteristic relations of regulatory linkage and behaviour. Here, we found simple expected patterns where a miRNA, miR157a-5p, was downregulated and its target

mRNA transcript encoding a plastid-localised protein (AT4G28660) in turn was upregulated, or vice versa (Figure 6a), but we also observed many miRNA targets that did not show any differential expression on the mRNA level, although corresponding miRNAs were differentially expressed. The effect of these miRNAs might be visible on the protein level due to inhibition of translation. If the target mRNA encodes a transcription factor, we should see the miRNA-dependent regulation in the expression of downstream targets of this transcription factor (Figure 6b), as reported before (Megraw *et al.*, 2016). As an example, the transcription factor AT4G36920 can act as an activator or repressor on its targets, and furthermore, the transcription factor can control other transcription factors or downstream targets like AT2G33380, which is a *CALEOSIN 3* transcript and important for stress responses (Sham *et al.*, 2015). In addition, the downstream transcription factors can also target other transcription factors or downstream targets, which increases the network complexity. Furthermore, this points to a sophisticated interaction between miRNAs and their targets, because miRNAs indirectly regulate genes encoding plastid proteins through the direct control of transcription factor mRNAs. Besides transcription factor mRNAs, many miRNAs are able to regulate transcripts of genes that do not encode transcription factors, but also these transcripts do not always show the expected behaviour. For instance, miR395c is predicted to control



**Figure 6.** Illustration of different network motifs which we observed in the miRNA–RNA target network in connection with relative changes of RNA levels between treatments. (a) Examples of expected regulations where a miRNA and its target mRNA exhibit inverted differential expression. (b) The interaction between a miRNA regulating the mRNA of a transcription factor, and the interactions of this transcription factor with its downstream target genes. (c) A downregulated miRNA which regulates four different targets. (d) An example of three miRNAs that regulate a single mRNA. The whole network can be accessed through the supporting Data S1 in GML format.

four different mRNA targets, including three transcripts encoding plastid-localised proteins (Figure 6c). Further, we found examples where several miRNAs are able to control the transcript of a single transcription factor (Figure 6d). These cases show possible interactions between miRNAs and their targets and suggest a wide range of direct and indirect impacts of miRNAs to regulate gene expression. Nevertheless, behavioural predictions are impossible without additional information on the exact mode of action of each miRNA and the magnitude of its influence.

## DISCUSSION

Until now, it is not known whether ncRNAs and sRNAs are regulated by retrograde signalling in response to NF treatment and how they contribute to the control of nuclear gene expression in response to plastid-derived signals. To better understand these biological processes, we combined sRNA sequencing with mRNA/lncRNA sequencing of *A. thaliana* WT seedlings and the two retrograde signalling mutants, *gun1* and *gun5*, to identify ncRNAs and mRNAs regulated by retrograde signals.

Generally, after NF treatment we detected nearly the same number of DEGs in all treated samples compared to the untreated WT. Further, we observed an overall tendency that more DEGs were downregulated than upregulated in response to NF treatment. In addition, we could observe an overrepresentation of DEGs encoding plastid-localised proteins in all three samples and detected more DEGs to be upregulated in the treated *gun* mutants compared to the treated WT.

Previous studies with different *gun* mutants were performed using *A. thaliana* microarrays lacking probes for ncRNAs (Strand *et al.*, 2003; Koussevitzky *et al.*, 2007; Woodson *et al.*, 2013). Koussevitzky *et al.* (2007) analysed changes in mRNA levels in WT (*Col-0*), *gun1* and *gun5* mutant seedlings grown on media with and without NF. About 43% of upregulated and 67% of downregulated DEGs in the present study overlap with those of Koussevitzky *et al.* (2007) in response to NF (Figure S9a,b). Generally, more downregulated DEGs and larger changes in the NF-treated *gun5* mutant than in the other mutant were identified in both studies. A good overlap of DEGs was found in both *gun* mutants. About 56% of the DEGs detected in *gun1-102* (*gun1* NF/WT NF) in our study were also detected in the treated *gun1-9* mutant by Koussevitzky *et al.* (2007) (Figure S9c), and about 50% of the DEGs identified in the treated *gun5* mutant (*gun5* NF/WT NF) in our study were also identified by Koussevitzky *et al.* (2007) (Figure S9d). However, in our data set we identified also 44% (*gun1* NF/WT NF) and 50% (*gun5* NF/WT NF) of DEGs that have not been shown to be controlled by *gun*-related retrograde signalling pathways before, which might be due to the differences between the two methods (RNA

sequencing versus microarrays) and the varying duration of the NF treatment between the studies (5 versus 4 days). Recently, another RNA sequencing study reported NF-responsive transcriptome changes in a different *gun1* mutant (*gun1-1*) (Richter *et al.*, 2020); 55% and 49% of the DEGs detected in the NF-treated *gun1* and *gun5* mutant compared to the NF-treated WT overlapped with DEGs in our study (Figure S9e,f). RNA sequencing was also performed in the *gun1-9* mutant grown in the presence of NF (Zhao *et al.*, 2019b) with an overlap of 65% of DEGs compared to our data (Figure S9g). Taken together, we observed a considerably high overlap with other transcriptome studies despite the differences between the studies regarding growth conditions, available mutants and analysis methods.

In our study, we observed an opposite regulation of differentially expressed plastid-encoded transcripts in both *gun* mutants, while the nuclear-encoded DEGs showed large overlap between the *gun1* and *gun5* mutants. Surprisingly, in response to NF all differentially expressed plastid-encoded transcripts were downregulated in the *gun1* mutant, whereas they were upregulated in the treated *gun5* mutant. These observations are in line with the model suggesting that plastid gene transcription is controlled by retrograde signalling networks, including sigma factors (SIG2 and SIG6) and plastid-encoded RNA polymerase (PEP), which might be crucial for proper plastid RNA transcription (Woodson *et al.*, 2013). It seems that *GUN1* activates PEP (Maruta *et al.*, 2015) and a perturbed PEP activation in the *gun1* mutant may prevent the upregulation of the plastid-encoded genes compared to WT upon NF treatment.

We identified an interesting lncRNA (AT4G13495) showing classical de-repression in both *gun* mutants (*gun* NF/WT NF) (Table 2). This lncRNA overlaps in sense direction with three different miRNA precursors (*MIR5026*, *MIR850* and *MIR863*) and all three miRNAs were differentially expressed in at least one treatment (WT NF/WT, *gun1* NF/WT NF and *gun5* NF/WT NF). We assume that all three miRNAs can be produced either from the three individual miRNA precursor transcripts or from the lncRNA. We did not find any predicted target for miR5026 according to the applied psRNATarget parameters. MiR850 was upregulated in the *gun5* mutant (*gun5* NF/WT NF), and two predicted cognate target RNAs, encoding a chloroplast RNA-binding protein (AT1G09340) and a threonine-tRNA ligase (AT2G04842), respectively, were upregulated as well. MiR863 targets the *SERRATE* transcript (AT2G27100), encoding an accessory protein essential for the miRNA biogenesis pathway, and thus may influence the regulation of several miRNAs (Meng *et al.*, 2012). MiR863 was upregulated in both treated *gun* mutants (*gun* NF/WT NF), but we did not detect significant changes of the *SERRATE* transcript in the two treated *gun* mutants.



Concerning the overall regulation of differentially expressed sRNAs belonging to different classes (miRNAs, nat-siRNAs and other sRNA producing loci), we detected more downregulated sRNAs in the treated WT (WT NF/WT), whereas both treated *gun* mutants exhibited a higher number of upregulated sRNAs (*gun* NF/WT NF). Principally, this suggests an increased sRNA processing in response to NF in both *gun* mutants, resembling the de-repression of nuclear-encoded genes, and we assume that these sRNAs might have an impact on retrograde-controlled nuclear gene expression. sRNAs are able to affect nuclear transcripts regulated by retrograde signals and they may regulate mRNA transcripts, affecting plastid-localised proteins. Among all sRNA classes, we observed in all treatments the highest number of differentially regulated sRNAs within the nat-siRNA class. Furthermore, all differentially regulated sRNAs have been associated to their corresponding putative differentially expressed RNA targets (Table S13) and we could detect high numbers of differentially regulated sRNAs. Besides an effect of sRNAs on their direct targets, we expect based on our network analyses a considerable indirect regulation by sRNAs through transcription factors (Data S1).

Interestingly, we found that miR169g-3p, a heat- and salt stress-responsive miRNA (Szyrajew *et al.*, 2017; Pegler *et al.*, 2019), is the most strongly downregulated miRNA in the treated WT (−151.6-fold; WT NF/WT), and the most strongly upregulated one in the treated *gun5* mutant (38.2-fold; *gun5* NF/WT NF). We did not find any predicted target for miR169g-3p according to our parameters.

Unexpectedly, after miRNA target prediction the expression of most of the targets was not anticorrelated to the expression changes of their cognate miRNA, leading us to conclude that miRNAs might not be involved in the expression of genes controlled by retrograde signalling pathways, or the expressional changes of miRNAs somehow balance transcriptional changes of their targets to maintain constant steady-state levels. Another possibility could be that they act as translational repressors and do not have a direct effect on the transcript abundance of their target RNAs. However, we predicted 20 miRNA targets coding for transcription factors and 22 targets encoding plastid-localised proteins to be targeted by 23 differentially regulated miRNAs. Thus, we assume that miRNAs may have important functions in the control of transcripts that code for regulatory proteins that are directly involved in transcriptional control and may contribute to the manifold changes of gene expression in response to retrograde signals. Further, nuclear transcripts that code for plastid-localised proteins are targets of miRNAs, suggesting that these specific miRNA–mRNA pairs can play an important role in the retrograde signalling pathway, and thus may contribute to the adjustment of plastidic and nuclear gene expression. One interesting case involves miR395b and

miR395c, which target the mRNA for the magnesium-chelatase subunit *GUN5* (AT5G13630). In the NF-treated WT, both miRNAs and the target mRNA are downregulated compared to the untreated control, whereas in the treated *gun5* mutant both miRNAs and the target are upregulated. Even though the expression of this miRNA–mRNA pair is not anticorrelated, the enhanced miRNA levels may balance an increased transcription rate of the target mRNA to keep physiologically relevant steady-state levels. Magnesium-chelatase is required in the chlorophyll biosynthesis pathway, where it catalyses the insertion of  $Mg^{2+}$  into protoporphyrin IX, and the *gun5* mutant is characterised by a single nucleotide substitution resulting in a defective magnesium-chelatase. In the WT, the *GUN5* transcript level decreases in response to NF-triggered retrograde signalling, whereas the transcript level in the *gun5* mutant remains high and cannot be efficiently downregulated by the increased miRNA levels. The seven detected classical anticorrelated miRNA–mRNA pairs point to regulatory functions of specific miRNAs in the retrograde signalling pathway, because we assume efficient miRNA-mediated target cleavage followed by a reduced mRNA steady-state level. In this category of anticorrelated pairs, we identified the mRNA for the transcription factor SPL10, representing a validated target of miR157a, suggesting miR157 acts in retrograde signalling by affecting the levels of a transcriptional regulator and its downstream targets. Another anticorrelated predicted miRNA–mRNA pair is miR398, targeting the transcript of the multidrug and toxic compound extrusion (MATE) efflux protein (AT2G04050). We found miR398 to be downregulated in the treated WT compared to the untreated control, and the target was slightly upregulated. This MATE efflux protein belongs to a huge class of membrane proteins located in the plasma membrane and the chloroplast envelope membrane (Wang *et al.*, 2016) that are able to bind cytotoxic compounds like primary and secondary metabolites, xenobiotic organic cations (Omote *et al.*, 2006) and toxic substances such as pollutants and herbicides (Diener *et al.*, 2001) to eliminate them from the cell (Liu *et al.*, 2016). NF-triggered downregulation of miR398 most likely causes elevated transcript levels and increased levels of the encoded plasma membrane-located MATE efflux protein. We speculate that the regulated MATE efflux protein might be involved in the extrusion of the applied herbicide NF or the extrusion of toxic compounds accumulating within the cell in response to NF treatment. Another interesting target encodes a plastid protein that appeared to be upregulated in the NF-treated *gun5* mutant (*gun5* NF/WT NF) by decreased levels of the cognate miRNAs. *PSB28* (AT4G28660), targeted by miR157a, encodes a protein that is part of the photosystem II reaction centre and is suggested to function in the biogenesis and assembly of chlorophyll-containing proteins (Mabbitt *et al.*, 2014). NF treatment usually leads to the

downregulation of *PhANGs* and thus should cause decreased expression levels of the *PSB28* mRNA. However, miR157a contributes to the downregulation of *PSB28* mRNA levels post-transcriptionally and seems to be controlled by retrograde signals, as indicated by the misregulation of miR157a in the *gun5* mutant.

Fang *et al.* (2018) identified miR395 and miR398 to be important in retrograde signalling triggered by tocopherols, and we confirmed both miRNAs applying NF as another trigger of retrograde signalling. Additionally, the transcript encoding the enzyme APS, which catalyses the initial step in PAP synthesis (Klein and Papenbrock, 2004; Pornsiriwong *et al.*, 2017) was identified to be targeted by miR395 (Liang *et al.*, 2010). We found miR395b to be downregulated in the treated WT (WT NF/WT) and upregulated in the treated *gun5* mutant (*gun5* NF/WT NF). In the treated WT, reduced miR395b levels lead to increased APS transcript levels, causing elevated PAP synthesis, which acts as retrograde inhibitor of XRN5 and should provoke elevated pri-miRNA and mature miRNA levels. Besides, Fang *et al.* (2018) detected the downregulation of miR398 in the WT after NF treatment, which is in line with our sRNA sequencing data. The COPPER/ZINC SUPEROXIDE DISMUTASE 2 (*CSD2*) was previously found to be a target of miR398 (Guan *et al.*, 2013). After heat stress, Fang *et al.* (2018) found miR398, PAP and tocopherol levels to be increased and *CSD2* levels to be decreased in the WT, and they hypothesised that tocopherols and PAP are required for miR398 biogenesis under heat stress. The *CSD2* mRNA escaped our miRNA target prediction due to a considerably high number of mismatches within the miRNA binding site, causing a score value that was above our cut-off value. Still, we identified this miRNA as differentially expressed supporting the previous study by Fang *et al.* (2018).

Besides differentially expressed miRNAs, we identified an even higher number of differentially regulated nat-siRNAs in the treated WT (WT NF/WT) and both *gun* mutants (*gun* NF/WT NF). Most of the overlapping transcripts encode nuclear or plastid proteins, suggesting that nat-siRNAs have a considerable impact on the control of *PhANGs* encoding plastid proteins. For most of the *cis*-NAT pairs we observed similar correlations between RNA transcript and sRNA expression levels. For example, the levels of two overlapping transcripts (AT1G05560 and AT1G05562) and the related *cis*-nat-siRNA were decreased in the treated WT (WT NF/WT) and increased in both *gun* mutants (*gun* NF/WT NF). The gene AT1G05562 encodes an antisense lncRNA and overlaps with the gene AT1G05560, which codes for a UDP-glucose transferase. Another interesting *cis*-nat-siRNA and one of the overlapping transcripts encoding a chlorophyll binding protein (AT1G29930) were downregulated in the treated WT (WT NF/WT) and upregulated in the treated *gun5* mutant (*gun5* NF/WT NF),

whereas levels of the other overlapping transcript, coding for a nuclear RNA polymerase (AT1G29940), remained unchanged in both treatments.

Here, we could demonstrate that NF treatment and subsequent retrograde signals lead to comprehensive changes in the steady-state levels of non-coding sRNAs comprising all known sRNA classes. The majority of the identified differentially expressed sRNAs belong to the *cis*- and *trans*-nat-siRNAs, followed by miRNAs, representing the second most abundant class. Thus, we postulate that mainly these two sRNA classes act as important regulators of gene expression in retrograde signalling. We also identified a considerably high number of so far unknown nuclear-encoded DEGs and thus add to the knowledge about genes that are controlled by retrograde signalling. Finally, we were able to identify promising sRNA-RNA target pairs that may act in the adjustment of plastidic and nuclear gene expression in retrograde signalling pathways.

## EXPERIMENTAL PROCEDURES

### Plant material and growth conditions

*Arabidopsis thaliana* WT (*Col-0*) and the retrograde signalling mutants *gun1-102* and *gun5-1* were used in this study. *Gun1-102* (SAIL\_290\_D09) harbours a transfer DNA insertion within the AT2G31400 gene locus resulting in a loss-of-function allele (Tadini *et al.*, 2016). *Gun5-1* is an EMS mutant harbouring a point mutation within the gene AT5G13630 causing an Ala/Val substitution at residue 990 (A990V) resulting in deficient magnesium-protoporphyrin IX synthesis (Mochizuki *et al.*, 2001). Surface-sterilised seeds were incubated on ½ MS agar plates containing 1.5% sucrose. For treatments with NF, seeds were incubated on the same medium supplemented with 5 µM norflurazon (Sigma-Aldrich, Taufkirchen, Germany). After vernalisation (2 days at 4°C in darkness) the seeds were grown for 4 days under continuous light (115 µmol photons m<sup>-2</sup> sec<sup>-1</sup>) at 22°C. Whole plants were harvested and immediately frozen in liquid nitrogen and stored at -80°C until RNA isolation. All control experiments and norflurazon treatments were performed in three biological replicates for each genotype.

### RNA isolation

The plant material was ground in liquid nitrogen and RNA isolation was performed using TRIzol reagent (Invitrogen) according to the manufacturer's protocol. RNA integrity was monitored by agarose gel electrophoresis and RNA concentration and purity were determined spectrophotometrically (260 nm/280 nm and 260 nm/230 nm absorbance ratios).

### sRNA purification

For sRNA sequencing 30 µg of total RNA were separated by 15% PAGE for 2 h at 120 V. The sRNA fractions with sizes ranging from 18 to 29 nucleotides were excised from the gel and eluted in 0.3 M NaCl overnight at 4°C with rotation. Remaining gel pieces were removed using a Spin-X centrifuge tube (Sigma-Aldrich) and 1 µl GlycoBlue (15 mg ml<sup>-1</sup>, Thermo Fisher), 25 µl sodium acetate (3 M, pH 5.0) and 625 µl ethanol were added to the 250 µl flow-through and samples were incubated for 4 h at -80°C. After centrifugation for 30 min with 17 000 g at 4°C the RNA pellet was

washed twice with 80% ethanol, dried and dissolved in nuclease-free water.

### Quantitative RT-PCR

Prior to cDNA synthesis, RNA samples were subjected to DNase I digestion (NEB, Ipswich, MA, USA) to remove residual genomic DNA. Total RNA (2 µg) was incubated at 37°C for 30 min together with DNase I (2 U; NEB). To inactivate the DNase I, 2.5 µl 50 mM EDTA was added and samples were incubated at 65°C for 10 min. The RNA was denatured for 5 min at 65°C in the presence of 100 pmol of an oligo-dT23VN oligonucleotide and 10 mM dNTPs and transferred to ice. Subsequently, cDNA synthesis was performed for 1 h at 42°C using M-MuLV reverse transcriptase (200 U; NEB) followed by heat inactivation at 80°C for 5 min. To monitor successful cDNA synthesis, we performed RT-PCR using gene-specific primers for the gene *UBI7* (AT3G53090) (Table S11).

For each qRT-PCR we used cDNA equivalent to 20 ng µl<sup>-1</sup> RNA with gene-specific primers and an EvaGreen qPCR mix. The samples were pre-heated for 2 min at 95°C and qRT-PCR cycling conditions were as follows: 12 sec at 95°C, 30 sec at 58°C and 15 sec at 72°C for 40 cycles. All qRT-PCRs were performed in three technical triplicates with the CFX Connect Real-Time PCR device (Bio-Rad, Feldkirchen, Germany). The Ct-values were used to calculate changes in gene expression by the 2<sup>-ΔΔCt</sup> method (Livak and Schmittgen, 2001). The values were normalised to the housekeeping gene *UBI1* (AT4G36800). Oligonucleotide sequences of all gene-specific primers are listed in Table S11.

### Stem-loop qRT-PCR

Stem-loop qRT-PCRs were used for sRNA quantification as described previously (Kramer, 2011). RNA from three independent biological replicates (300 ng) was used for cDNA synthesis. RNA was denatured for 5 min at 65°C together with 100 pmol of stem-loop oligonucleotides (Table S11) and 10 mM dNTPs. cDNA synthesis was performed for 5 min at 25°C and 20 min at 42°C using M-MuLV reverse transcriptase (200 U; NEB), followed by heat inactivation at 80°C for 5 min. RT-PCR for the gene *UBI7* (AT3G53090) served as control (Table S11).

### RNA sequencing

For the generation of mRNA libraries, including poly(A)-tailed lncRNAs, 10 µg total RNA from each sample was vacuum-dried in the presence of RNAsable (Sigma-Aldrich). The libraries were prepared using the Next Ultra RNA Library Prep Kit (NEB) by Novogene (China). The samples were sequenced strand-specifically as 150 bp paired-end reads on a HiSeq-PE150 platform with at least 15 million read pairs per library.

sRNA libraries for each RNA sample were generated twice following two slightly modified protocols. The first set of libraries was generated from 5 µg total RNA with the NEBNext Multiplex Small RNA Library Prep Set for Illumina according to the manufacturer's instructions and 1 h of 3' adapter ligation. The second set of sRNA libraries was prepared from purified sRNAs obtained from 30 µg of total RNA using the same kit as described above performing 3' adapter ligation for 18 h. For both libraries, excessive non-ligated 3' adapters were made inaccessible by converting them into dsRNA by hybridisation of complementary oligonucleotides. 5' adapter ligation was carried out at 25°C for 1.5 h, reverse transcription was performed by using the ProtoScript II reverse transcriptase and libraries were amplified by 12 PCR cycles. The PCR products were separated by 6% PAGE for 2 h at

60 V. The cDNA library fractions with a size ranging from 138 to 150 nucleotides were excised from the gel and eluted overnight. The sRNA libraries were sequenced as 50 bp single-end reads on an Illumina HiSeq1500 sequencer with approximately 10 million reads per library.

### Analysis of mRNA and lncRNA

The mRNA and lncRNA sequencing results were analysed with the open source and web-based platform GALAXY (Afgan *et al.*, 2016). The FASTQ raw sequences were trimmed with the tool Trimmomatic to remove adapter sequences with their default parameters (Bolger *et al.*, 2014). Tophat (Kim *et al.*, 2013) was used to map the reads against the *A. thaliana* reference genome (<https://www.arabidopsis.org/>, release: TAIR10) with a maximum intron length parameter of 3000 nt. The transcripts were annotated in Araport11 (<https://apps.araport.org/thalemine/dataCategories.do>); we considered annotated ncRNAs longer than 200 bp as lncRNAs. Differential expression of transcripts was analysed by Cuffdiff (Trapnell *et al.*, 2010) to normalise the sequencing depth of each library and to calculate FPKM values. The FDR was used as a statistic indicator to exclude type I errors or rather false positives. Transcripts having FDR ≤ 0.05 and log<sub>2</sub>(FC) ≤ -1 and ≥ +1 were considered as DEGs. The package pheatmap (<https://cran.r-project.org/web/packages/pheatmap/pheatmap.pdf>) was used to generate hierarchically (UPGMA) clustered heatmaps of differentially expressed RNAs (Kolde, 2019).

### Gene ontology terms

GO enrichment terms were analysed using the DAVID Bioinformatics Resources 6.8 (<https://david.ncifcrf.gov>) with default parameters (Huang *et al.*, 2009a,b) and results were visualised with the R package 'ggpubr' (Wickham, 2016).

### Analysis of sRNA

NEBNext Kit adapter sequences were clipped from the sequencing reads using a custom script within GALAXY that identifies Illumina adapter sequences using a seed sequence of 10 nt. After adapter clipping FASTQ files of the raw reads with a length of 18–26 nt were loaded into the CLC Genomics Workbench 11.0.1 (Qiagen, Hilden, Germany) for further analyses. The ShortStack analysis package was used for advanced analysis of the sRNA sequences (Axtell, 2013b). The FASTQ files of the six biological replicates derived from each treatment were first mapped against the *A. thaliana* TAIR10 reference genome (<https://www.arabidopsis.org/>, release: TAIR10). The merged alignments were mapped against a file covering all *A. thaliana* mature miRNAs (<http://www.mirbase.org/>) and a second file comprising different RNA classes, namely nat-siRNAs, ta-siRNAs, phasiRNAs and lncRNAs. A nat-siRNA database (Table S1) was generated from previously annotated NAT pairs (Jin *et al.*, 2008; Zhang *et al.*, 2012; Yuan *et al.*, 2015), phasiRNAs were taken from Howell *et al.* (Howell *et al.*, 2007) and lncRNAs were downloaded from Araport11 (<https://apps.araport.org/thalemine/dataCategories.do>, release: Araport 11 Annotation). After mapping to the respective references, the individual raw reads for each replicate were used for normalisation and differential expression analysis based on a calculation with DESeq2 (Love *et al.*, 2014). The sRNAs were filtered by fold changes between ≤ -2 and ≥ +2. The significance of the differentially expressed sRNAs was evaluated with FDR ≤ 0.05. MiRNA target RNAs were identified using the psRNATarget (Dai *et al.*, 2018) prediction V2 tool (<http://plantgrn.noble.org/psRNATarget/>) from protein-coding and non-coding transcripts present in Araport11. An expectation value of less than 2.5 was considered as a cut-off

for true miRNA targets, where mRNAs harbouring a lower number of mismatches to the reverse and complementary miRNA obtain lower score values.

### Network analysis

Using the sRNA and mRNA sequencing data together with the miRNA target prediction, we assembled an interaction network of miRNAs and their putative targets. For network analysis, we used the python package networkX (Hagberg *et al.*, 2008). We further investigated the miRNA targets that were predicted as described above using the psRNATarget tool to identify all miRNA targets that encode transcription factors. For this, we compared all miRNA targets with a reference database, containing *A. thaliana* transcription factors, which was generated using publicly available data (<http://atrm.cbi.pku.edu.cn>). Furthermore, this reference database was extended by incorporating available information about whether the transcription factors act as activators or repressors of gene expression together with available information about the individual target genes of the transcription factors (<https://agris-knowledgebase.org>). The data obtained from the RNA sequencing experiments were then used to generate a network of miRNAs and their targets, differentiating between miRNA targets encoding transcription factors and targets encoding other proteins. For network analyses connected to measurements, only RNAs with a FDR less than 0.05 were considered, unless indicated otherwise. Network analyses were performed to determine relationships between miRNAs and their targets with special focus on miRNA targets encoding transcription factors. In these network analyses we considered the impact of miRNAs on the transcripts coding for transcription factors and the impact of the expression of transcription factor mRNAs on downstream genes, regulated by these transcription factors. If the change of a miRNA results in an expected change of an mRNA coding for a transcription factor, or at least downstream genes show expected transcriptional changes, we classified this behaviour as 'expected'. For example, if miRNA expression was reduced and the target mRNA encoding a transcription factor was upregulated, the transcription factor acted as an activator and downstream genes of this transcription factor were consequently also upregulated, this would be considered as 'expected' behaviour. Furthermore, the scatter plots presented for all differentially expressed miRNAs ( $FDR \leq 0.05$ ) were subdivided into two categories: plots depicting relations of miRNA and their direct target transcripts (direct) and plots depicting indirect relations comprising miRNAs that control mRNAs encoding transcription factors and their downstream target genes (indirect).

### ACCESSION NUMBERS

ATG accession numbers: GUN1 (AT2G31400), GUN5 (AT5G13630). The raw Illumina sRNA and mRNA/lncRNA sequencing data are deposited in the NCBI SRA database with the ID PRJNA557616.

### ACKNOWLEDGEMENTS

This work was supported by the German Research Foundation (SFB-TRR 175, grants to WF project C03 and EK project D03). We thank Martin Simon for technical support and advice regarding sRNA library preparations and Oguz Top for helpful comments on the manuscript. Open access funding enabled and organized by Projekt DEAL.

### AUTHOR CONTRIBUTIONS

WF and MAA designed the research; KH performed the research with help from MAA and BT; KH, MAA, BT and WF analysed the data; network analysis was performed by SOA, MK and EK; and KH, MAA, MK, SOA and WF wrote the paper. All authors read and approved the final manuscript.

### CONFLICT OF INTEREST

The authors declare no conflict of interest.

### SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

**Appendix S1.** Figure and table captions.

**Figure S1.** Effect of NF on the plants and sRNA size distribution.

**Figure S2.** Distribution of differentially expressed *cis*- and *trans*-nat-siRNA in response to NF.

**Figure S3.** Distribution of differentially expressed sRNAs derived from lncRNAs in response to NF.

**Figure S4.** Validation of transcript levels by qRT-PCR.

**Figure S5.** Distribution of differentially expressed ncRNAs in the untreated and treated samples.

**Figure S6.** GO term enrichment analysis.

**Figure S7.** Validation of sRNAs and their associated transcripts.

**Figure S8.** Distributions of miRNA–RNA target interaction numbers.

**Figure S9.** Comparison of differentially expressed mRNAs from Koussevitzky *et al.* (2007), Richter *et al.* (2020), Zhao *et al.* (2019) and our study.

**Table S1.** Reference sequences for the *cis*- and *trans*-NAT pairs.

**Table S2.** sRNA sequencing and mapping results for each independent biological replicate.

**Table S3.** Lists of all significant differentially expressed sRNA classes.

**Table S4.** RNA mapping results after the analysis with Tophat for each independent biological replicate.

**Table S5.** Overview of all significant differentially expressed mRNAs and other RNA classes.

**Table S6.** Accession numbers of all significant differentially expressed mRNAs and other RNA classes.

**Table S7.** List of the individual clusters of the mRNA heatmap.

**Table S8.** GO term enrichment analysis for significant DEGs.

**Table S9.** Lists of miRNA target prediction.

**Table S10.** Lists of *cis*- and *trans*-nat-siRNAs and their target correlation.

**Table S11.** Sequences of oligonucleotides used in this study.

**Table S12.** Distributions of miRNA–RNA target interaction numbers.

**Table S13.** Overview of all significant differentially expressed (DE) sRNAs and their corresponding differentially expressed mRNA target.

**Data S1.** MiRNA–RNA target network in GML format that can be accessed with the free software 'Gephi' (Bastian *et al.*, 2009).



## REFERENCES

- Afgan, E., Baker, D., van den Beek, M. *et al.* (2016) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.* **44**, W3–W10.
- Axtell, M.J. (2013a) Classification and comparison of small RNAs from plants. *Annu. Rev. Plant Biol.* **64**, 137–159.
- Axtell, M.J. (2013b) ShortStack: comprehensive annotation and quantification of small RNA genes. *RNA*, **19**, 740–751.
- Ball, L., Accotto, G.P., Bechtold, U. *et al.* (2004) Evidence for a direct link between glutathione biosynthesis and stress defense gene expression in Arabidopsis. *Plant Cell*, **16**, 2448–2462.
- Bannister, A.J. and Kouzarides, T. (2011) Regulation of chromatin by histone modifications. *Cell Res.* **21**, 381–395.
- Bartel, D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281–297.
- Bastian, M., Heymann, S. and Jacomy, M. (2009) Gephi: an open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media*.
- Bolger, A.M., Lohse, M. and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.
- Borsani, O., Zhu, J., Verslues, P.E., Sunkar, R. and Zhu, J.K. (2005) Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. *Cell*, **123**, 1279–1291.
- Breitenbach, J., Zhu, C. and Sandmann, G. (2001) Bleaching herbicide norflurazon inhibits phytoene desaturase by competition with the cofactors. *J. Agric. Food Chem.* **49**, 5270–5272.
- Chan, K.X., Phua, S.Y., Crisp, P., McQuinn, R. and Pogson, B.J. (2016) Learning the languages of the chloroplast: Retrograde signaling and beyond. *Annu. Rev. Plant Biol.* **67**, 25–53.
- Chen, X. (2009) Small RNAs and their roles in plant development. *Annu. Rev. Cell Dev. Biol.* **25**, 21–44.
- Dai, X., Zhuang, Z. and Zhao, P.X. (2018) psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic Acids Res.* **46**, W49–W54.
- Diener, A.C., Gaxiola, R.A. and Fink, G.R. (2001) Arabidopsis ALF5, a multidrug efflux transporter gene family member, confers resistance to toxins. *Plant Cell*, **13**, 1625–1638.
- Dinger, M.E., Pang, K.C., Mercer, T.R., Crowe, M.L., Grimmond, S.M. and Mattick, J.S. (2009) NRED: a database of long noncoding RNA expression. *Nucleic Acids Res.* **37**, D122–D126.
- Fang, X., Zhao, G., Zhang, S., Li, Y., Gu, H., Li, Y., Zhao, Q. and Qi, Y. (2018) Chloroplast-to-nucleus signaling regulates microRNA biogenesis in Arabidopsis. *Dev Cell*, **48**, 371–382.
- Fettke, J., Nunes-Nesi, A., Fernie, A.R. and Steup, M. (2011) Identification of a novel heteroglycan-interacting protein, HIP 1.3, from *Arabidopsis thaliana*. *J. Plant Physiol.* **168**, 1415–1425.
- Franco-Zorrilla, J.M., Valli, A., Todesco, M., Mateos, I., Puga, M.I., Rubio-Somoza, I., Leyva, A., Weigel, D., Garcia, J.A. and Paz-Ares, J. (2007) Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat. Genet.* **39**, 1033–1037.
- Goksoyr, J. (1967) Evolution of eucaryotic cells. *Nature*, **214**, 1161.
- Gray, J.C., Sullivan, J.A., Wang, J.H., Jerome, C.A. and Maclean, D. (2003) Coordination of plastid and nuclear gene expression. *Philos. Trans. R Soc. Lond B Biol. Sci.* **358**, 135–145.
- Guan, Q., Lu, X., Zeng, H., Zhang, Y. and Zhu, J. (2013) Heat stress induction of miR398 triggers a regulatory loop that is critical for thermotolerance in Arabidopsis. *Plant J.* **74**, 840–851.
- Gutierrez-Nava, M.D.L., Gillmor, C.S., Jimenez, L.F., Guevara-Garcia, A. and Leon, P. (2004) CHLOROPLAST BIOGENESIS genes act cell and noncell autonomously in early chloroplast development. *Plant Physiol.* **135**, 471–482.
- Hagberg, A., Schult, D. and Swart, P. (2008). Exploring Network Structure, Dynamics, and Function using NetworkX. Proceedings of the 7th Python in Science conference (SciPy 2008). 11–15.
- Heo, J.B. and Sung, S. (2011) Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science*, **331**, 76–79.
- Holoch, D. and Moazed, D. (2015) RNA-mediated epigenetic regulation of gene expression. *Nat. Rev. Genet.* **16**, 71–84.
- Hong, Z., Zhang, Z., Olson, J.M. and Verma, D.P. (2001) A novel UDP-glucose transferase is part of the callose synthase complex and interacts with phragmoplastin at the forming cell plate. *Plant Cell*, **13**, 769–779.
- Howell, M.D., Fahlgren, N., Chapman, E.J., Cumbie, J.S., Sullivan, C.M., Givan, S.A., Kasschau, K.D. and Carrington, J.C. (2007) Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in Arabidopsis reveals dependency on miRNA- and tasiRNA-directed targeting. *Plant Cell*, **19**, 926–942.
- Huang, C.Y., Wang, H., Hu, P., Hamby, R. and Jin, H. (2019) Small RNAs - Big players in plant-microbe interactions. *Cell Host Microbe*, **26**, 173–182.
- Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009a) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13.
- Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009b) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57.
- Jin, H., Vacic, V., Girke, T., Lonardi, S. and Zhu, J.K. (2008) Small RNAs and the regulation of cis-natural antisense transcripts in Arabidopsis. *BMC Mol. Biol.* **9**(6). <https://doi.org/10.1186/1471-2199-9-6>
- Kakizaki, T., Matsumura, H., Nakayama, K., Che, F.S., Terauchi, R. and Inaba, T. (2009) Coordination of plastid protein import and nuclear gene expression by plastid-to-nucleus retrograde signaling. *Plant Physiol.* **151**, 1339–1353.
- Khraiwesh, B., Arif, M.A., Seumel, G.I., Ossowski, S., Weigel, D., Reski, R. and Frank, W. (2010) Transcriptional control of gene expression by microRNAs. *Cell*, **140**, 111–122.
- Kim, C. and Apel, K. (2013) 102-mediated and EXECUTER-dependent retrograde plastid-to-nucleus signaling in norflurazon-treated seedlings of *Arabidopsis thaliana*. *Mol. Plant*, **6**, 1580–1591.
- Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36.
- Kim, V.N. (2005) Small RNAs: classification, biogenesis, and function. *Mol. Cells*, **19**, 1–15.
- Klein, M. and Papenbrock, J. (2004) The multi-protein family of Arabidopsis sulphotransferases and their relatives in other plant species. *J. Exp. Bot.* **55**, 1809–1820.
- Kleine, T. and Leister, D. (2016) Retrograde signaling: organelles go networking. *Biochim. Biophys. Acta*, **1857**, 1313–1325.
- Kolde, R. (2019) Pretty Heatmaps. 1.0. 12 ed. <https://cran.r-project.org/web/packages/heatmap/heatmap.pdf>
- Koussevitzky, S., Nott, A., Mockler, T.C., Hong, F., Sachetto-Martins, G., Surpin, M., Lim, J., Mittler, R. and Chory, J. (2007) Signals from chloroplasts converge to regulate nuclear gene expression. *Science*, **316**, 715–719.
- Kramer, M.F. (2011) Stem-loop RT-qPCR for miRNAs. *Curr. Protoc. Mol. Biol.* **95**(1), 15.10.1–15.10.15. <https://doi.org/10.1002/0471142727.mb1510s95>
- Lapidot, M. and Pilpel, Y. (2006) Genome-wide natural antisense transcription: coupling its regulation to its different regulatory mechanisms. *EMBO Rep.* **7**, 1216–1222.
- Larkin, R.M., Alonso, J.M., Ecker, J.R. and Chory, J. (2003) GUN4, a regulator of chlorophyll synthesis and intracellular signaling. *Science*, **299**, 902–906.
- Liang, G., Yang, F. and Yu, D. (2010) MicroRNA395 mediates regulation of sulfate accumulation and allocation in *Arabidopsis thaliana*. *Plant J.* **62**, 1046–1057.
- Liu, J., Li, Y., Wang, W., Gai, J. and Li, Y. (2016) Genome-wide analysis of MATE transporters and expression patterns of a subgroup of MATE genes in response to aluminum toxicity in soybean. *BMC Genom.* **17**, 223.
- Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, **25**, 402–408.
- Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550.
- Lu, Y. (2016) Identification and roles of photosystem II assembly, stability, and repair factors in Arabidopsis. *Front. Plant Sci.* **7**, 168.
- Ma, X., Shao, C., Jin, Y., Wang, H. and Meng, Y. (2014) Long non-coding RNAs: a novel endogenous source for the generation of Dicer-like 1-dependent small RNAs in *Arabidopsis thaliana*. *RNA Biol.* **11**, 373–390.
- Mabbitt, P.D., Wilbanks, S.M. and Eaton-Rye, J.J. (2014) Structure and function of the hydrophilic Photosystem II assembly proteins: Psb27, Psb28 and Ycf48. *Plant Physiol. Biochem.* **81**, 96–107.

- Maruta, T., Miyazaki, N., Nosaka, R. *et al.* (2015) A gain-of-function mutation of plastidic invertase alters nuclear gene expression with sucrose treatment partially via GENOMES UNCOUPLED1-mediated signaling. *New Phytol.* **206**, 1013–1023.
- Megraw, M., Cumbie, J.S., Ivanchenko, M.G. and Filichkin, S.A. (2016) Small genetic circuits and microRNAs: big players in polymerase II transcriptional control in plants. *Plant Cell*, **28**, 286–303.
- Meister, G. and Tuschl, T. (2004) Mechanisms of gene silencing by double-stranded RNA. *Nature*, **431**, 343–349.
- Meng, Y., Shao, C., Ma, X., Wang, H. and Chen, M. (2012) Expression-based functional investigation of the organ-specific microRNAs in Arabidopsis. *PLoS One*, **7**, e50870.
- Meskauskiene, R., Nater, M., Goslings, D., Kessler, F., op den Camp, R. and Apel, K. (2001) FLU: a negative regulator of chlorophyll biosynthesis in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. USA*, **98**, 12826–12831.
- Meyers, B.C., Axtell, M.J., Bartel, B. *et al.* (2008) Criteria for annotation of plant MicroRNAs. *Plant Cell*, **20**, 3186–3190.
- Mochizuki, N., Brusslan, J.A., Larkin, R., Nagatani, A. and Chory, J. (2001) Arabidopsis genomes uncoupled 5 (GUN5) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. *Proc. Natl. Acad. Sci. USA*, **98**, 2053–2058.
- Molla-Morales, A., Sarmiento-Manus, R., Robles, P., Quesada, V., Perez-Perez, J.M., Gonzalez-Bayon, R., Hannah, M.A., Willmitzer, L., Ponce, M.R. and Micol, J.L. (2011) Analysis of ven3 and ven6 reticulate mutants reveals the importance of arginine biosynthesis in Arabidopsis leaf development. *Plant J.* **65**, 335–345.
- Omote, H., Hiasa, M., Matsumoto, T., Otsuka, M. and Moriyama, Y. (2006) The MATE proteins as fundamental transporters of metabolic and xenobiotic organic cations. *Trends Pharmacol. Sci.* **27**, 587–593.
- Park, W., Li, J., Song, R., Messing, J. and Chen, X. (2002) CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. *Curr. Biol.* **12**, 1484–1495.
- Pegler, J.L., Oultram, J.M.J., Grof, C.P.L. and Eamens, A.L. (2019) Profiling the abiotic stress responsive microRNA landscape of Arabidopsis thaliana. *Plants (Basel)*, **8**, 58.
- Pornsiriwong, W., Estavillo, G.M., Chan, K.X. *et al.* (2017) A chloroplast retrograde signal, 3'-phosphoadenosine 5'-phosphate, acts as a secondary messenger in abscisic acid signaling in stomatal closure and germination. *Elife*, **6**, e23361.
- Richter, A.S., Tohge, T., Fernie, A.R. and Grimm, B. (2020) The genomes uncoupled-dependent signalling pathway coordinates plastid biogenesis with the synthesis of anthocyanins. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **375**, 20190403.
- Rossel, J.B., Walter, P.B., Hendrickson, L., Chow, W.S., Poole, A., Mullineaux, P.M. and Pogson, B.J. (2006) A mutation affecting ASCORBATE PEROXIDASE 2 gene expression reveals a link between responses to high light and drought tolerance. *Plant Cell Environ.* **29**, 269–281.
- Saini, G., Meskauskiene, R., Pijacka, W., Roszak, P., Sjogren, L.L., Clarke, A.K., Straus, M. and Apel, K. (2011) 'happy on norflurazon' (hon) mutations implicate perturbation of plastid homeostasis with activating stress acclimatization and changing nuclear gene expression in norflurazon-treated seedlings. *Plant J.* **65**, 690–702.
- Sham, A., Moustafa, K., Al-Ameri, S., Al-Azzawi, A., Iratni, R. and Abuqamar, S. (2015) Identification of Arabidopsis candidate genes in response to biotic and abiotic stresses using comparative microarrays. *PLoS One*, **10**, e0125666.
- Strand, A., Asami, T., Alonso, J., Ecker, J.R. and Chory, J. (2003) Chloroplast to nucleus communication triggered by accumulation of Mg-protoporphyrinIX. *Nature*, **421**, 79–83.
- Susek, R.E., Ausubel, F.M. and Chory, J. (1993) Signal transduction mutants of Arabidopsis uncouple nuclear CAB and RBCS gene expression from chloroplast development. *Cell*, **74**, 787–799.
- Swiezewski, S., Liu, F., Magusin, A. and Dean, C. (2009) Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. *Nature*, **462**, 799–802.
- Szyrajew, K., Bielewicz, D., Dolata, J., Wojcik, A.M., Nowak, K., Szczygiel-Sommer, A., Szwejkowska-Kulinska, Z., Jarmolowski, A. and Gaj, M.D. (2017) MicroRNAs are intensively regulated during induction of somatic embryogenesis in Arabidopsis. *Front. Plant Sci.*, **8**, 18.
- Tadini, L., Pesaresi, P., Kleine, T. *et al.* (2016) GUN1 controls accumulation of the plastid ribosomal protein S1 at the protein level and interacts with proteins involved in plastid protein homeostasis. *Plant Physiol.* **170**, 1817–1830.
- Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J. and Pachter, L. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515.
- Voinnet, O. (2009) Origin, biogenesis, and activity of plant microRNAs. *Cell*, **136**, 669–687.
- Wang, H.V. and Chekanova, J.A. (2017) Long Noncoding RNAs in Plants. *Adv. Exp. Med. Biol.* **1008**, 133–154.
- Wang, L., Bei, X., Gao, J., Li, Y., Yan, Y. and Hu, Y. (2016) The similar and different evolutionary trends of MATE family occurred between rice and Arabidopsis thaliana. *BMC Plant Biol.* **16**, 207.
- Wickham, H. (2016) *ggplot2 – Elegant Graphics for Data Analysis*. New York: Springer-Verlag.
- Wierzbicki, A.T., Haag, J.R. and Pikaard, C.S. (2008) Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell*, **135**, 635–648.
- Woodson, J.D., Perez-Ruiz, J.M. and Chory, J. (2011) Heme synthesis by plastid ferrochelatase I regulates nuclear gene expression in plants. *Curr. Biol.* **21**, 897–903.
- Woodson, J.D., Perez-Ruiz, J.M., Schmitz, R.J., Ecker, J.R. and Chory, J. (2013) Sigma factor-mediated plastid retrograde signals control nuclear gene expression. *Plant J.* **73**, 1–13.
- Yuan, C., Wang, J., Harrison, A.P., Meng, X., Chen, D. and Chen, M. (2015) Genome-wide view of natural antisense transcripts in Arabidopsis thaliana. *DNA Res.* **22**, 233–243.
- Zhang, X., Xia, J., Lii, Y.E. *et al.* (2012) Genome-wide analysis of plant natural RNAs reveals insights into their distribution, biogenesis and function. *Genome Biol.* **13**, R20.
- Zhao, C., Wang, Y., Chan, K.X. *et al.* (2019a) Evolution of chloroplast retrograde signaling facilitates green plant adaptation to land. *Proc. Natl. Acad. Sci. USA*, **116**, 5015–5020.
- Zhao, X., Huang, J. and Chory, J. (2019b) GUN1 interacts with MORF2 to regulate plastid RNA editing during retrograde signaling. *Proc. Natl. Acad. Sci. USA*, **116**, 10162–10167.
- Zimorski, V., Ku, C., Martin, W.F. and Gould, S.B. (2014) Endosymbiotic theory for organelle origins. *Curr. Opin. Microbiol.* **22**, 38–48.

## Publication II: Identification of cold stress regulated small RNAs in *Arabidopsis thaliana*

Bhavika Tiwari<sup>1</sup>, Kristin Habermann<sup>1</sup>, M. Asif Arif<sup>1</sup>, Heinrich Lukas Weil<sup>2</sup>, Timo Mühlhaus<sup>2</sup>, Wolfgang Frank<sup>1\*</sup>

1 Plant Molecular Cell Biology, Department Biology I, Ludwig-Maximilians-Universität München, LMU Biocenter, 82152 Planegg-Martinsried, Germany

2 Computational Systems Biology Division, TU Kaiserslautern, Paul-Ehrlich-Straße 23, 67663, Kaiserslautern, Germany

<https://doi.org/10.1186/s12870-020-02511-3>

### Abstract

**Background:** Cold stress causes dynamic changes in gene expression that are partially caused by small non-coding RNAs since they regulate protein coding transcripts and act in epigenetic gene silencing pathways. Thus, a detailed analysis of transcriptional changes of small RNAs (sRNAs) belonging to all known sRNA classes such as microRNAs (miRNA) and small interfering RNA (siRNAs) in response to cold contributes to an understanding of cold-related transcriptome changes.

**Result:** We subjected *A. thaliana* plants to cold acclimation conditions (4 °C) and analyzed the sRNA transcriptomes after 3 h, 6 h and 2 d. We found 93 cold responsive differentially expressed miRNAs and only 14 of these were previously shown to be cold responsive. We performed miRNA target prediction for all differentially expressed miRNAs and a GO analysis revealed the overrepresentation of miRNA-targeted transcripts that code for proteins acting in transcriptional regulation. We also identified a large number of differentially expressed *cis*- and *trans*-nat-siRNAs, as well as sRNAs that are derived from long non-coding RNAs. By combining the results of sRNA and mRNA profiling with miRNA target predictions and publicly available information on transcription

factors, we reconstructed a cold-specific, miRNA and transcription factor dependent gene regulatory network. We verified the validity of links in the network by testing its ability to predict target gene expression under cold acclimation.

**Conclusion:** In *A. thaliana*, miRNAs and sRNAs derived from *cis*- and *trans*-NAT gene pairs and sRNAs derived from lncRNAs play an important role in regulating gene expression in cold acclimation conditions. This study provides a fundamental database to deepen our knowledge and understanding of regulatory networks in cold acclimation.



RESEARCH ARTICLE

Open Access



# Identification of small RNAs during cold acclimation in *Arabidopsis thaliana*

Bhavika Tiwari<sup>1</sup>, Kristin Habermann<sup>1</sup>, M. Asif Arif<sup>1</sup>, Heinrich Lukas Weil<sup>2</sup>, Antoni Garcia-Molina<sup>3</sup>, Tatjana Kleine<sup>3</sup>, Timo Mühlhaus<sup>2</sup> and Wolfgang Frank<sup>1\*</sup>

## Abstract

**Background:** Cold stress causes dynamic changes in gene expression that are partially caused by small non-coding RNAs since they regulate protein coding transcripts and act in epigenetic gene silencing pathways. Thus, a detailed analysis of transcriptional changes of small RNAs (sRNAs) belonging to all known sRNA classes such as microRNAs (miRNA) and small interfering RNA (siRNAs) in response to cold contributes to an understanding of cold-related transcriptome changes.

**Result:** We subjected *A. thaliana* plants to cold acclimation conditions (4 °C) and analyzed the sRNA transcriptomes after 3 h, 6 h and 2 d. We found 93 cold responsive differentially expressed miRNAs and only 14 of these were previously shown to be cold responsive. We performed miRNA target prediction for all differentially expressed miRNAs and a GO analysis revealed the overrepresentation of miRNA-targeted transcripts that code for proteins acting in transcriptional regulation. We also identified a large number of differentially expressed *cis*- and *trans*-nat-siRNAs, as well as sRNAs that are derived from long non-coding RNAs. By combining the results of sRNA and mRNA profiling with miRNA target predictions and publicly available information on transcription factors, we reconstructed a cold-specific, miRNA and transcription factor dependent gene regulatory network. We verified the validity of links in the network by testing its ability to predict target gene expression under cold acclimation.

**Conclusion:** In *A. thaliana*, miRNAs and sRNAs derived from *cis*- and *trans*-NAT gene pairs and sRNAs derived from lncRNAs play an important role in regulating gene expression in cold acclimation conditions. This study provides a fundamental database to deepen our knowledge and understanding of regulatory networks in cold acclimation.

**Keywords:** *Arabidopsis thaliana*, Cold acclimation, Small non-coding RNA, Gene regulation, RNA sequencing, miRNA-transcription factor network

## Background

Plants are severely affected by dynamic and extreme climatic conditions. Changes in temperature is one of the most critical factors for plants to exhibit flourishing growth and low temperature stress globally influences the development of plants and restricts their spatial distribution affecting the total agricultural productivity [1].

Although most plant species have evolved a certain degree of cold tolerance, deviations from the optimal conditions lead to restructuring at the gene level enabling the plant to cope with the environmental fluctuations [2].

Plant cells perceive cold stress by detecting reduced cell membrane fluidity that triggers specific signaling cascades [3] to induce the expression of cold responsive genes [4]. Currently, the best characterized pathway is the C-repeat binding factor (CBF)-dependent signaling pathway in which OPEN STOMATA 1 (OST1)/SNF1-related protein kinase 2 (SnRK2.6/SnRK2E) is released from type 2C protein phosphatase (PP2Cs) in response

\* Correspondence: [wolfgang.frank@lmu.de](mailto:wolfgang.frank@lmu.de)

<sup>1</sup>Department of Biology I, Plant Molecular Cell Biology, Ludwig-Maximilians-Universität München, LMU Biocenter, Großhaderner Str. 2-4, 82152 Planegg-Martinsried, Germany  
Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

to elevated abscisic acid (ABA) [5] levels to activate the upstream transcription factor (TF) inducer of CBF expression (ICE1) by phosphorylation [6]. ICE1 further induces the expression of several CBF/ dehydration responsive element binding factors (DREB) TFs that bind to the cold response sensitive TFs/dehydration responsive elements (CRT/DRE) promoter elements of cold-responsive (*COR*) genes, which act in the adaptation to low temperature conditions [7, 8]. Another ABA-dependent pathway that controls *COR* gene expression is mediated through the binding of bZIP TFs known as ABRE-binding factors (ABFs) to ABA-responsive promoter elements [9, 10]. Furthermore, studies have shown that DREB/CBF can physically interact with ABFs to express ABA responsive genes [11]. The CRT/DRE and ABRE regions are present in many cold-inducible genes and indicate a tight link between the ABA-dependent pathway and the ICE-CBF-COR pathway [10].

In addition to the TF mediated transcriptional control, epigenetic modifications control the gene expression in cold stress mainly by chromatin remodeling altering the accessibility of chromatin for the transcription machinery [12, 13]. Besides the transcriptional control, gene regulation involves regulatory processes at the post-transcriptional and post-translational level [14]. An important post-transcriptional control of gene expression is mediated by non-coding RNAs (ncRNAs) that cannot be translated into functional proteins. ncRNAs are classified into long non-coding RNAs (lncRNAs) that contribute to the control of gene expression involving transcriptional and post-transcriptional pathways [15] and sRNAs binding to reverse complementary target RNAs to confer target RNA cleavage or translational inhibition [16] or they interfere with transcription via epigenetic mechanisms such as RNA-directed DNA methylation (RdDM) [17].

lncRNAs are longer than 200 nt and possess 5' capping and 3' polyadenylation similar to mRNAs [18–20]. lncRNAs exert their function by different modes of action, for instance lncRNAs restrain the accessibility of regulatory proteins to nucleic acids by serving as decoys [21]. Another mechanism is presented by the well characterized lncRNA Induced by Phosphate Starvation1 (*IPS1*), that acts as a non-cleavable competitor for the *Phosphate 2* (*PHO2*) mRNA that is targeted by miR399 for degradation [22]. lncRNAs also cause epigenetic alterations such as histone modifications as identified in the vernalization process where prolonged cold stress leads to epigenetic silencing of the Flowering locus C (*FLC*) that controls flowering time [23, 24]. Here, the lncRNA cold induced long antisense intragenic RNA (*COOLAIR*) interacts with a polycomb repressive complex (PRC2) and subsequently causes histone methylation and silencing of the *FLC* locus. lncRNAs also assist in de novo methylation of DNA cytosine residues and cause transcriptional silencing of genes by RdDM [25, 26].

Small RNAs (sRNA) are 21–24 nt in size and efficiently regulate mRNA transcript levels, translation and also mediate epigenetic silencing [27]. The two main sRNA classes are microRNA (miRNAs) that are processed from single stranded precursors forming a partially double-stranded hairpin structure and small interfering RNAs (siRNAs) that are generated from double-stranded RNA precursors. miRNA biogenesis occurs in a multi-step fashion starting with the transcription of nuclear encoded *MIR* genes by RNA polymerase II to produce a 5' capped and polyA-tailed primary miRNA transcript (pri-miRNA) [28]. The dicing complex containing Dicer-like1 (DCL1) and its accessory proteins Hyponastic Leaves 1 (HYL1) and Serrate (SE) excise a miRNA duplex from the double stranded hairpin structure that is translocated to the cytoplasm by the exportin Hasty (HST). The mature miRNA is loaded into an argonaute protein within the RNA-induced silencing complex to mediate the cleavage of target mRNAs via reverse complementary binding of the miRNA [29].

Plant miRNAs play important roles in a wide range of biological processes including development and stress adaptation [30]. To uncover the stress-regulated miRNA repertoire, sRNA libraries were generated from plants subjected to diverse stress conditions and analyzed by RNA sequencing approaches [31–34]. Previous studies in *A. thaliana* identified members of the miR171 family to be upregulated by low as well as elevated temperatures [35] targeting *SCARECROW-LIKE6-III* (*SCL6-III*) and *SCL6-IV* that belong to the *GRAS* family of TFs [36, 37]. MiR408 was recognized to be induced by cold and other abiotic stresses. It regulates transcripts encoding phytochrome family proteins (cupredoxin, plantacyanin and uclacyanin) which act as electron transfer shuttles between proteins [38] and transcripts of phytophenol oxidases called Laccases [39] which are known to oxidize flavonoids during seed development and environmental stress [40]. These are essential to maintain cell wall functions and are important to regulate biological pathways necessary for abiotic stress responses [41]. Recent investigations validated miR394 and its target *LEAF CURLING RESPONSIVENESS* (*LCR*) to regulate leaf development [42, 43] and to be involved in an ABA-dependent manner in responses to cold, salt and drought stress [44, 45]. In *A. thaliana*, miR397 was shown to positively regulate cold tolerance via the CBF-dependent signaling pathway and overexpression of *MIR397a* caused increased CBF transcript levels leading to induction of cold responsive *COR* genes [46].

In contrast to miRNAs, siRNAs are generated from dsRNA molecules and are sub-classified based on their specific biogenesis pathways. *Trans*-acting siRNAs (ta-siRNAs) are endogenous plant-specific small RNAs that are capable of acting in *trans* and have the potential to

repress distinct mRNA transcripts. The production of ta-siRNAs is triggered by miRNA-mediated cleavage of primary *TAS* transcripts to generate 21 nt ta-siRNAs in a phased manner [47, 48]. Ta-siRNAs have been shown to regulate plant development [49]. Recent studies suggest their role in environmental stress adaptation, for example, 14 hypoxia-responsive ta-siRNAs have been identified in *A. thaliana* that are processed from *TAS1a*, *b*, *c*, *TAS2* and *TAS3a* precursors [50]. The expression of a *TAS1*-derived ta-siRNA and its target transcript *heat-induced TAS1 target (HTT4)* were shown to be regulated by temperature shifts [51]. Furthermore, the generation of *TAS4*-derived ta-siRNAs was shown to be triggered by miR828 under phosphate deficiency [52].

Another subset of siRNAs are natural antisense transcript derived short interfering RNAs (nat-siRNAs) which are produced from overlapping regions of RNA polymerase II derived antisense transcripts [53]. The NATs can be classified into two types depending on the genomic location of the overlapping transcripts. Either both transcripts are encoded on opposite DNA strands within the same genomic region to produce overlapping transcripts (*cis*-NATs) or both transcripts derive from separate genomic regions (*trans*-NATs), but are able to pair with each other. A high salinity responsive nat-siRNA was first identified in *A. thaliana* where the constitutively expressed gene transcript *delta-pyrroline-5-carboxylate dehydrogenase (P5CDH)* and the salt induced gene transcript *Similar to Radicle Induced Cell Death One 5 (SRO5)* encoded on opposing strands of an overlapping genomic region form a dsRNA and DCL2 processes a distinct 24 nt nat-siRNA from the dsRNA region. The generated nat-siRNA cleaves the *P5CDH* transcript and suppresses proline degradation thereby inducing salinity tolerance [54]. In addition to nat-siRNAs produced from *cis*-NATs, *trans*-NATs can be generated when antisense-mediated pairing of transcripts occurs that are derived from non-overlapping genes [55]. The formation of these dsRNAs takes place in diverse *trans*-combinations i.e. between long non-coding RNAs, protein coding transcripts, homologous pseudogenes and transposable elements (TE) [56, 57]. For example, the class of *trans*-NATs that are produced from pseudogenes can regulate their homologous protein encoding transcripts levels [58].

A large number of TE-derived siRNAs were observed in *Decreased DNA methylation 1 (DDM1)* mutants of *A. thaliana* and are referred to as epigenetically activated siRNAs (ea-siRNAs). These siRNAs are produced from transposon-encoded transcripts that are cleaved in a miRNA-dependent manner and become converted into dsRNAs that are further processed by DCL4 into 21 nt ea-siRNAs. These ea-siRNAs were shown to be mainly required for silencing of TE by targeting their intrinsic

transcripts whereas a subset of these siRNAs also targets protein coding mRNAs to reduce their expression levels [59]. In addition, similar to *MIR* precursors some TE-derived transcripts can form a stem loop structure from which siRNAs can be processed [60]. TE also encode lncRNAs and there is rising evidence that environmental factors lead to altered chromatin organization and the expression of lncRNAs that may have functions in the adaptation to altered environmental conditions and can even be inherited. A study in *A. thaliana* reports on a TE-derived TE-lincRNA1195 that was shown to be involved in the ABA response and to contribute to abiotic stress adaptation [61].

In our study we have used RNA sequencing to uncover the cold responsive non-coding RNA repertoire in *A. thaliana* and to study their role in the regulation of various target RNAs. We sequenced mRNAs and sRNAs libraries from *A. thaliana* plants subjected to cold acclimation conditions for 3 h, 6 h and 2 d and analyzed putative correlations between differentially expressed sRNAs and their protein coding targets. To gain additional insight into the cold-responsive interconnection of miRNA-regulated direct targets and indirect targets that are regulated by TFs, we generated a gene regulatory network (GRN) using information on miRNA-targets and publicly available TF-related database the generated network allows to identify connectivities and regulatory impacts of miRNAs under cold acclimation.

## Results

### Altered expression of sRNAs during cold acclimation in *A. thaliana*

To analyze cold-responsive changes in the sRNA repertoire we subjected *A. thaliana* seedlings to 4 °C cold treatment for 3 h, 6 h and 2 d time points. Previous studies related to cold acclimation observed a rapid inhibition of photosynthetic machinery when shifted from normal temperatures to 4 °C [62]. In addition, studies revealed that abundant cold-responsive genes were differentially expressed at early time points i.e. 3 h and 6 h as well as at later time points i.e. 48 h [33, 34, 62]. Thus, in order to study the sRNAs that could possibly regulate these cold-altered genes, the 3 h, 6 h and 2 d time points were chosen for RNA sequencing analyses. The RNA of treated and control samples were used to perform transcriptome profiling yielding a minimum of 7 million reads per library. The sRNA reads were mapped to the *A. thaliana* reference genome and in all samples on average about 10% reads mapped to miRNA loci, 10% to *trans*- and 2% to *cis*-nat-siRNA loci, 4% reads mapped to lncRNAs, 3% to ta-siRNA producing regions and 0.3% to pha-siRNAs (Additional file 1: Table S2). Only about 1% of the total reads mapped to loci encoding the most abundant RNAs such as ribosomal RNA, snoRNA, tRNA

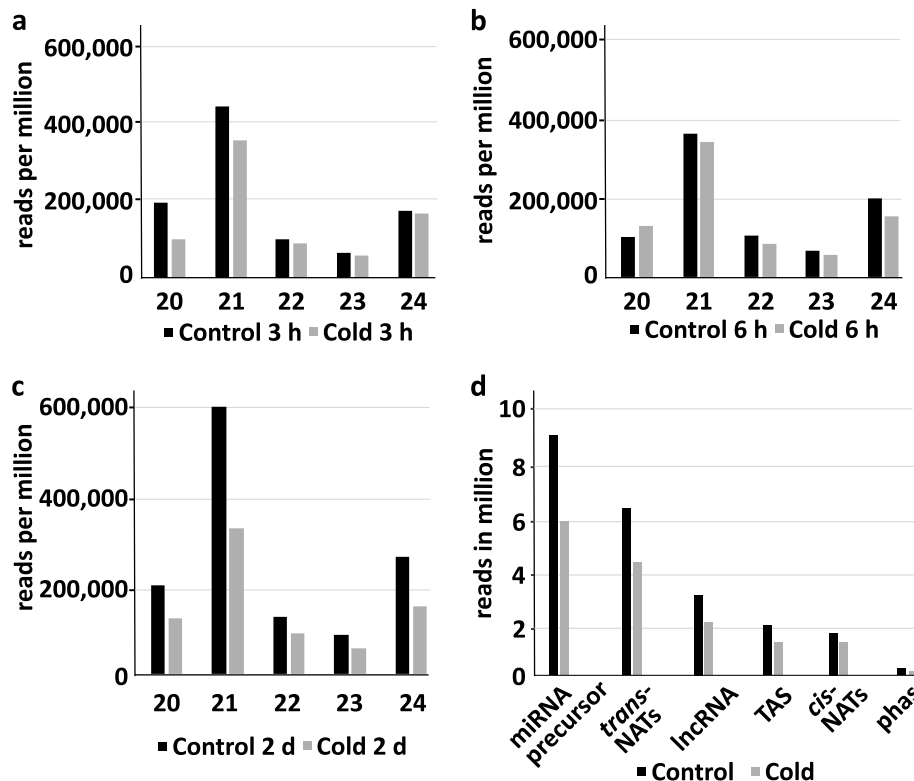
and snRNA which indicates a good quality of the sRNA libraries. The remaining proportion of reads mostly mapped to other RNA classes such as TE and repeat associated regions which are known to be involved in epigenetic pathways.

The size distribution of sRNAs ranging from 21 to 24 nt showed two distinct peaks at 21 nt indicating an enrichment of miRNAs, nat-siRNAs and ta-siRNA and at 24 nt corresponding to sRNAs derived from repetitive/intergenic RNAs, inverted repeats and TE (Fig. 1a, b, c, Additional file 1: Table S3). We observed an overall reduction of sRNAs in response to cold acclimation as compared to the control. The distribution of sRNA reads mapping to different sRNA producing loci including miRNAs, nat-siRNAs, ta-siRNAs, phasiRNAs and sRNAs produced from lncRNAs indicated that miRNAs and *trans*-nat-siRNAs are the two major sRNA classes detected in our data set (Fig. 1d). To identify differentially expressed (DE) sRNAs between cold treated samples and the respective untreated controls (fold change  $\geq 2$  &  $\leq -2$  and a Benjamini-Hochberg corrected  $p$ -value  $\leq 0.05$ ), the relative expression of mature miRNAs and siRNAs was calculated on the basis of the number of normalized reads. Over the analyzed time course cold stress mainly affected sRNAs produced from *trans*- and

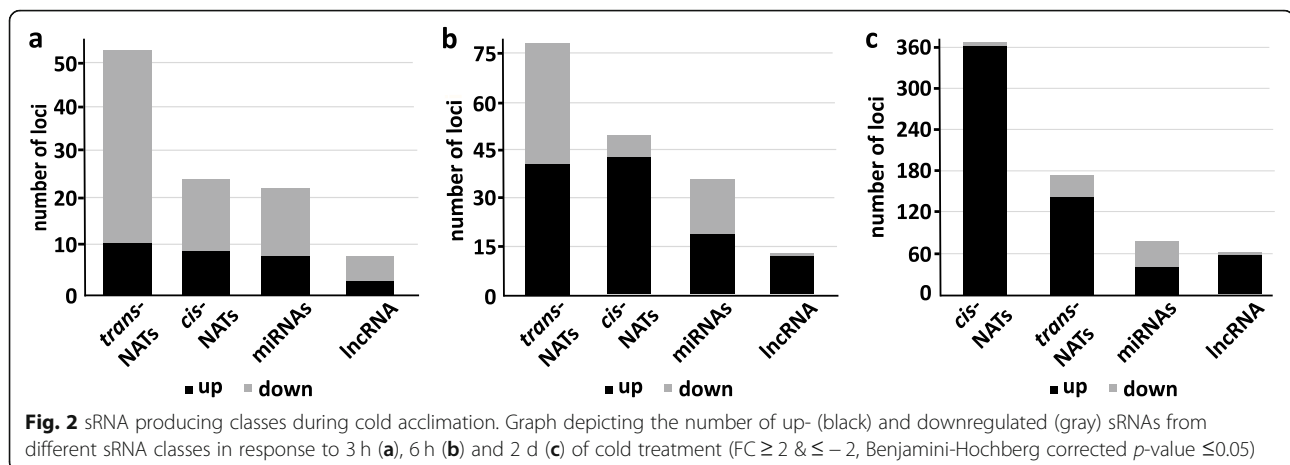
*cis*-NATs-pairs followed by the class of miRNAs and sRNAs derived from lncRNA (Fig. 2a, b, c). Moreover, we observed an increasing number of up- and downregulated sRNAs from all sRNA classes during the time course reaching the highest numbers after 2 d of the cold treatment (Fig. 2c). To evaluate the reliability of the sRNA sequencing results, we performed stem-loop qRT-PCRs for selected sRNAs belonging to all analyzed sRNA classes to validate and confirm their expressional changes during the time course of cold treatment (Fig. 3). miR162a-3p, miR3434-5p, *cis*-nat-siRNA produced from *AT3G05870-AT3G05880* transcripts, a *trans*-nat-siRNA generated from *AT1G10522-AT5G53905* transcripts and a sRNA derived from lncRNA *AT5G04445* were found to be induced over the course of cold treatment confirming our sRNA sequencing results.

#### Expression profiling of cold acclimation responsive miRNAs

The sRNA sequencing method allows to distinguish between individual miRNAs with even a single nucleotide difference. After precise read mapping, sequence reads were analyzed to identify differentially regulated miRNAs (FC  $\geq 2$  &  $\leq -2$ , Benjamini-Hochberg corrected  $p$ -value  $\leq 0.05$ ) (Table 2, Additional file 2: Table S4). We



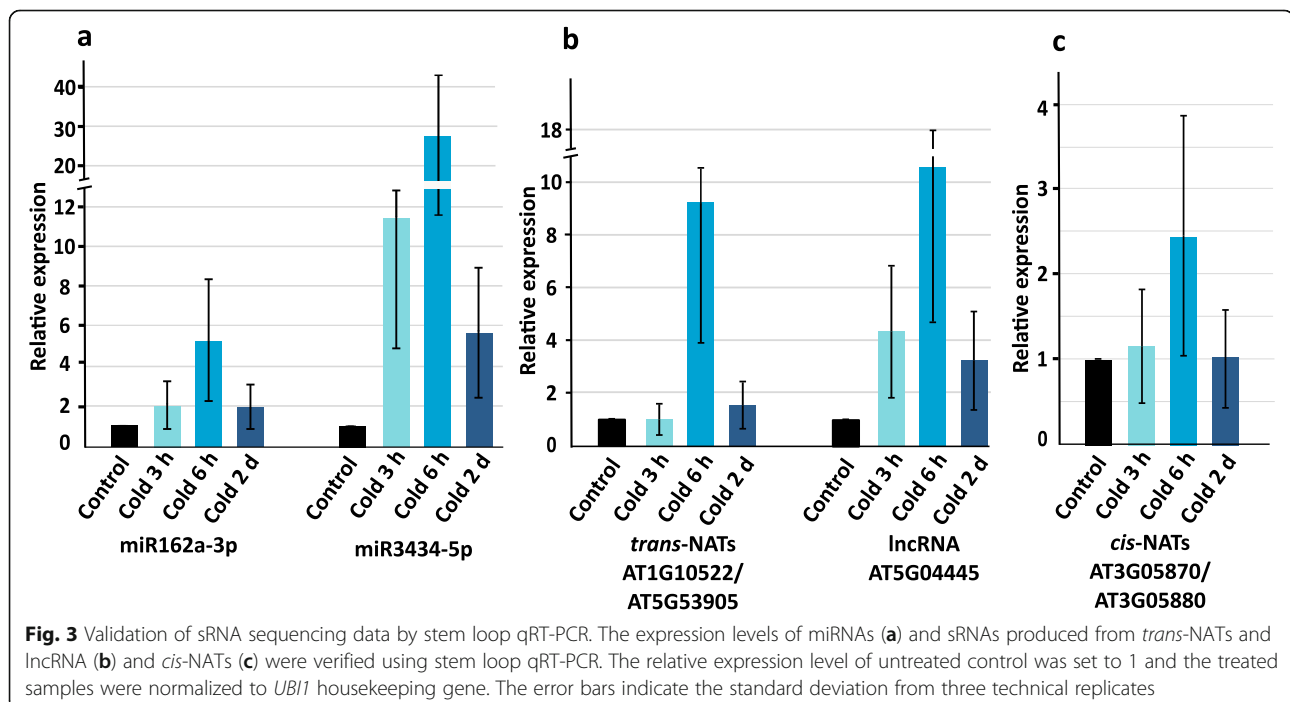
**Fig. 1** sRNA size distribution. Graphs depicting the size distribution of mapped sRNAs ranging from 20 to 24 nt in response to cold treatment and in the respective untreated controls after 3 h (a), 6 h (b) and 2 d (c) (represented in reads per million). Average trimmed sRNA reads per million mapping to different classes of RNAs in control and cold acclimated samples (d)



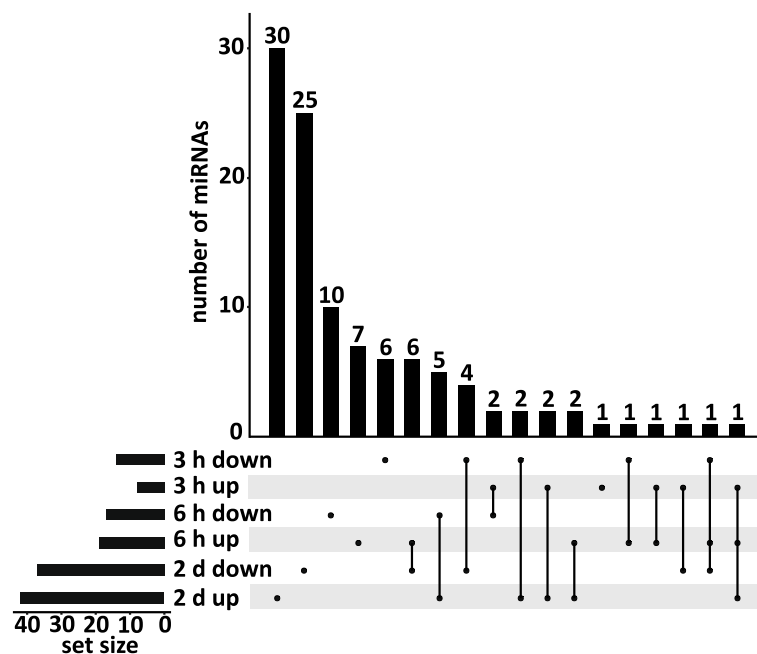
observed a general trend in all the samples that around 10% of the detected miRNAs possessed very high normalized read counts ( $> 1000$  reads per sample), about 50% showed moderate expression ( $< 1000$  and  $> 20$  normalized reads), 11% showed reduced read counts ( $< 20$  and  $> 5$  normalized reads) and 27% showed very low expression ( $< 5$  normalized reads) (Additional file 2: Table S5). In response to cold treatment we observed 22 miRNAs (8 up and 14 down) that were DE after 3 h, 36 mature DE miRNAs (19 up and 17 down) after 6 h and 79 DE mature miRNAs (42 up and 37 down) after 2 d. We found miRNAs showing differential expression at specific time points as well as miRNAs with differential expression at two or all three time points. Two DE miRNAs were found throughout the course of cold

treatment, 13 DE miRNAs were detected at 6 h and 2 d, 8 DE miRNAs were common after 3 h and 2 d, and 4 DE miRNAs were found at the 3 h and 6 h time point. We also observed 7, 17 and 55 DE miRNAs that were specifically regulated at the 3 h, 6 h, and 2 d time points (Fig. 4). We detected an increasing number of DE individual miRNAs over the time course of cold treatment suggesting that alterations in miRNA levels seem to be an important step during cold acclimation.

In recent years, 22 miRNA families were identified to be conserved between *A. thaliana*, *Oryza sativa* and *Populus trichocarpa* [63–65] and some of them were shown to have important roles in abiotic stress adaptation since they predominantly regulate targets encoding TFs or enzymes that promote tolerance to stresses [66–68]. Out of these







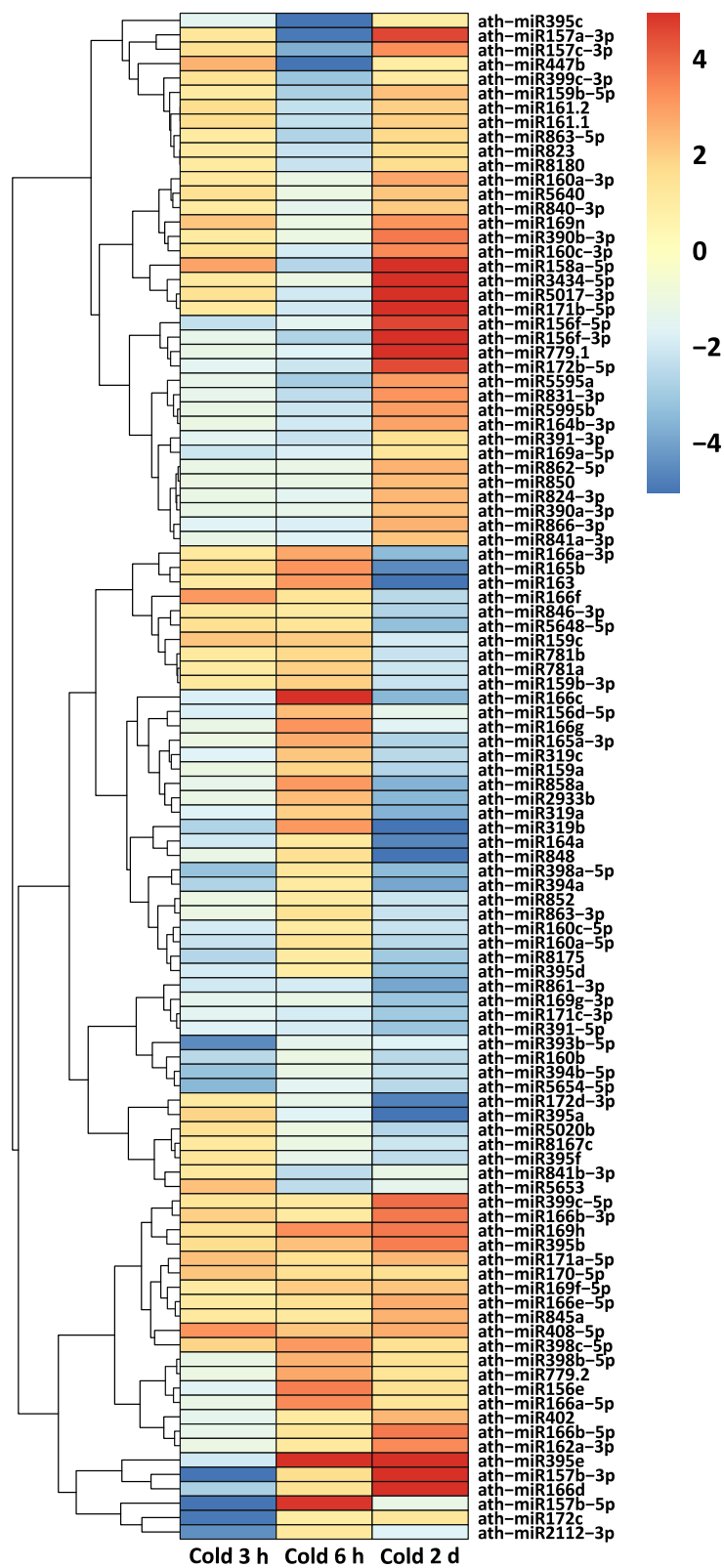
**Fig. 4** UpSet plot depicting the number of DE miRNAs. The plot depicts the global comparison of up- and downregulated DE miRNAs after 3 h, 6 h and 2 d of cold treatment ( $FC \geq 2$  &  $\leq -2$ , Benjamini-Hochberg corrected  $p$ -value  $\leq 0.05$ )

22 miRNA families, we detected individual members of 16 families to be differentially expressed corresponding to 15, 20 and 43 DE mature miRNAs at 3 h, 6 h and 2 d, respectively (Fig. 5, Additional file 2: Table S6). In total, we found 107 non-redundant mature miRNAs to be differentially expressed throughout the course of cold treatment and 36 mature miRNAs out of these belonging to 9 conserved miRNA families have been known to be cold regulated in other plant species (Additional file 2: Table S6) [35, 69, 70]. Out of 107 miRNAs, 14 have been previously known to be cold responsive in *A. thaliana* and our study shows similarity in the induction or repression pattern of these miRNAs compared to other cold stress related studies [35, 71, 72]. The remaining 93 DE mature miRNAs that belong to 55 miRNA families have not been reported before to be cold-regulated in *A. thaliana* (Additional file 2: Table S6). We identified several miRNAs with a varying expression pattern i.e. up- and downregulation at different time points. For example, miR156f-5p and miR157b-3p were downregulated at 3 h and upregulated at 2 d, miR166f was upregulated at 3 h and downregulated at the 2 d time point, miR447b and miR5653 were upregulated at 3 h, but downregulated at 6 h time point whereas miR157b-5p was downregulated at 3 h and upregulated at 6 h. Similarly, 12 miRNAs showed inconsistent regulation at 6 h and 2 d, whereas we observed consistent upregulation of miR408-5p, miR395e, miR159c, miR169h and downregulation of miR160a-

5p, miR160b, miR398a-5p, miR8175, miR319b in at least two time points. This indicates that the regulatory pattern of a miRNA can vary at different time points of cold treatment and the steady-state level of mature miRNAs depends on the physiological need of plants subjected to stress conditions.

#### Differentially expressed miRNA targets

Since miRNAs and mRNA/lncRNA were sequenced from the same RNA samples we were able to compare changes in miRNA expression with the changes of their cognate targets. To identify the targets of miRNAs that were found to be differentially expressed during the time course of cold treatment we have used the psRNAtarget prediction tool with a stringent expectation cut-off of 2.5 and allowed miRNA accessibility to its mRNA target with a maximum energy to unpair the target site of 25 [73]. Applying these stringent parameters, the prediction tool revealed putative targets for 93 DE miRNAs out of 107. The target prediction for the 93 non-redundant DE miRNAs identified 338 mRNAs and 14 non-coding RNAs as putative targets (Additional file 3: Table S7, S8). The 18 DE miRNAs at 3 h (5 up- and 13 downregulated) can target 96 non-redundant mRNAs and 3 non-coding transcripts. The 33 DE miRNAs at 6 h (18 up- and 15 downregulated) can target 173 non-redundant mRNAs and 3 non-coding RNA targets and the 69 DE miRNAs after 2 d (34 up- and 35 downregulated) are able to target 267 non-redundant mRNAs and 12 non-coding RNA targets (Additional file 3: Table S7, S8). To



**Fig. 5** Hierarchically clustered heatmap depicting miRNAs differentially expressed in at least one of the analyzed time points in response to cold treatment

analyze how the regulation of these targets correlates with the expression of miRNAs, we used our mRNA and lncRNA transcriptome sequencing data generated from the identical RNA pools as the sRNA data set for the 3 h, 6 h and 2 d cold treatments and their respective untreated controls (Additional file 4: Table S9, S10). We used the mRNA/lncRNA transcriptome data to examine the expression levels of all 338 transcripts targeted by the 93 differentially regulated miRNAs in order to correlate the target transcript expression to the expression of their cognate miRNAs (Additional file 3: Table S7). In frequent cases we observed that one transcript can be targeted by various isoforms of a miRNA family, but in a few cases target transcripts can also be cleaved by different miRNAs that are unrelated in sequence. In general, we considered all individual DE miRNAs and their cognate protein-coding targets (mRNAs) as miRNA:mRNA pairs and identified 111, 246 and 376 of these pairs for the 3 h, 6 h and 2 d time points of cold treatment, respectively (Additional file 3: Table S7). For each time point we classified the miRNA:mRNA target pairs into different subgroups according to the correlation of their expression with the expression of their cognate miRNA. These miRNA:mRNA target pair subgroups were classified as inversely correlated when they show an anticorrelation of mRNA and miRNA expression, showing same tendency of expression when the miRNA and its target are either upregulated or downregulated, and the miRNA is regulated, but the target remains unchanged or undetected (Table 1). We observed 2, 12 and 27 anticorrelated miRNA:mRNA target pairs at 3 h, 6 h and 2 d, respectively, with a total number of 39 non-redundant anticorrelated miRNA:mRNA target pairs pointing to a role of these miRNAs in controlling the transcriptome upon cold treatment (Additional file 3: Table S7). Apart from the mRNA targets, the target prediction tool also identified 14 putative non-coding RNA targets of DE miRNAs, but the expression levels of ncRNA target transcripts was less than 5 reads or they were not differentially expressed.

**Table 1** Number of putative miRNA:mRNA target pairs and their relative expression pattern after 3 h, 6 h and 2 d of cold treatments

miRNA:mRNA pairs	3 h	6 h	2 d
↑ ↓	2	12	27
↑ ↑	0	9	15
↓ ↓	1	1	12
↑ — or ↓ —	7	15	32
↑ ○ or ↓ ○	29	137	99

The first arrow corresponds to miRNA regulation and the second to the regulation of its target mRNA transcripts and the arrows represent the correlation expression as follows: ↑ = upregulated, ↓ = downregulated, — = unchanged, ○ = undetected.

On the basis of Araport (Version 11; <https://araport.org/>) annotation, we observed 54 targets of DE miRNAs from all the four subgroups to be consistently present at all the time points (Additional file 5: Table S11). These mainly encode TFs and DNA binding domain containing proteins and include MYB domain containing proteins, nuclear factor Y subunit genes, heat shock TFs (HsFs), TCP domain proteins and Squamosa promoter binding (SPLs) proteins. We also examined the functions of the miRNA targets that were specific for each time point. Specifically, at 6 h time point we found several PPR proteins that are known to be important for RNA maturation in various organelles, TPR encoding genes required in plant signaling and organellar import and genes encoding membrane multi-antimicrobial extrusion [22] efflux proteins that act in the transport of xenobiotic compounds. At the 2 d time point we found abundant transcripts coding for factors involved in transcriptional regulation and protein phosphorylation that control intracellular signaling in response to stress. Taken together, we found a remarkable overrepresentation of genes encoding transcription factors, proteins associated with transcriptional regulation, and proteins involved in RNA processing and translational control.

We found 39 miRNAs and their putative targets showing an inverse correlation, for example, after 3 h of cold treatment we noticed a strong downregulation of miR172c (FC = -4.86) and an upregulation of its predicted target TARGET OF EARLY ACTIVATION TAGGED (EAT, FC = 2.18) which is known to be reduced in *A. thaliana ice1* mutants [33]. In addition, EAT also showed increased expression levels in roots and leaves at 4 °C in *A. thaliana* [74]. After 6 h of cold treatment we observed downregulation of miR395c (FC = -19.27) and a concomitant upregulation of its target transcript encoding the magnesium-chelatase subunit H which presents the GUN5 gene (FC = 2.18) that was shown to be an important component of plastid to nucleus signal transduction. Another miRNA, miR5595a showed reduced expression levels (FC = -2.88) whereas its target encoding a methyl esterase 9 was upregulated (FC = 3.58) and is known to be a plant core environmental stress responsive gene (PCESR) [75]. Additionally, after 2 d of cold treatment, we observed three isoforms of miR319 to be downregulated and an upregulation of one of their target transcript encoding a TCP2 TF (FC = 2.56). A previous study revealed an upregulation of the TCP2 transcript after shifting *A. thaliana* plants to cold conditions with 100 μE light conditions, but not in dark conditions and it was speculated that light-dependent signals derived from the chloroplast at low temperature are important for increased TCP2 levels that might be important for the control of photosynthesis related genes [76, 77]. After 2 d of cold treatment we also detected downregulation of miR159 isoforms (FC = -2.53) resulting in elevated levels of

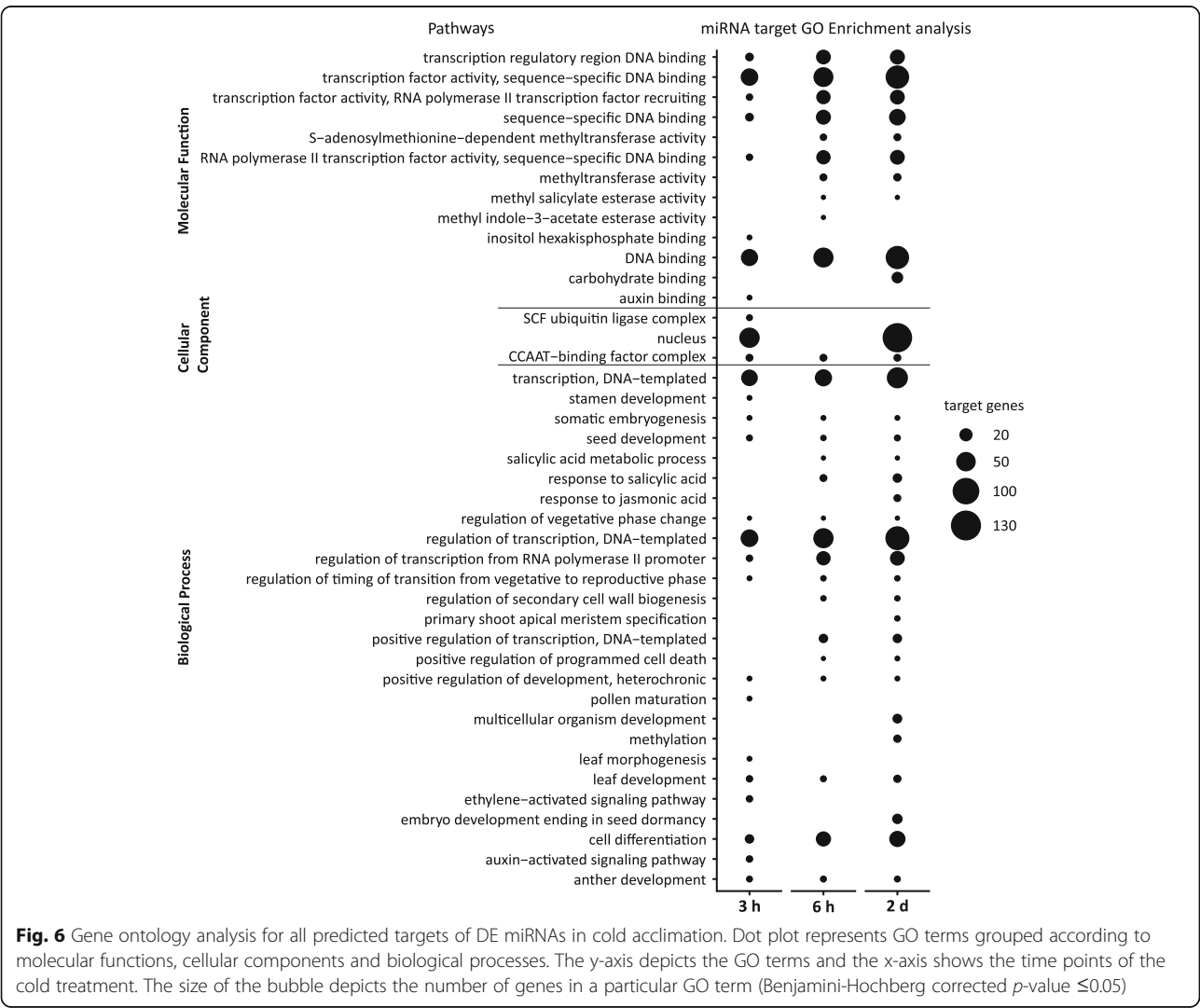


one of their target transcripts *Translocase Inner Membrane Subunit 44 (TIM44)-related* encoding a subunit of the mitochondrial inner membrane translocase complex subunit (FC = 2.80).

### Gene ontology analysis of predicted miRNA targets

To obtain information about the possible role of DE cold responsive miRNAs and their targets, we performed gene ontology (GO) analysis of all putative targets using the David bioinformatics tool [78]. Based on the three categories; biological processes, cellular component and molecular function, we observed an enrichment of GO terms for all three time points with Benjamini-Hochberg corrected *p*-values obtained from Fisher's test (Fig. 6, Additional file 6: Table S12). At the 3 h time point the significant biological processes included regulation of transcription (47), transcription (41), cell differentiation (12), ethylene-activated signaling pathway (7) and auxin-activated signaling pathway (7) indicating a major impact

of miRNAs on an early response of genes that code for proteins mainly acting in signaling and gene transcription. Concerning the category cellular component, we identified the highest number of targets associated with the nucleus (63) which nicely correlates with the overrepresentation of TFs before. Furthermore, in the category molecular functions, the TF activity, sequence-specific DNA binding (46), DNA binding (44) and auxin binding functions were most significant also pointing to an overrepresentation of transcripts that code for regulatory proteins and factors involved in gene transcription. For the 6 h time point significant biological processes with the highest number of genes included regulation of transcription (62 target genes), response to salicylic acid (8), regulation of secondary cell wall biogenesis (5) and positive regulation of programmed cell death. We also found S-adenosylmethionine-dependent methyltransferase activity (7) to be significantly enriched in the molecular function category. Similar to 3 h time point, we



**Fig. 6** Gene ontology analysis for all predicted targets of DE miRNAs in cold acclimation. Dot plot represents GO terms grouped according to molecular functions, cellular components and biological processes. The y-axis depicts the GO terms and the x-axis shows the time points of the cold treatment. The size of the bubble depicts the number of genes in a particular GO term (Benjamini-Hochberg corrected *p*-value ≤ 0.05)

observed an enrichment of transcription related genes at the 6 h time point. Along with these, the overrepresentation of methyltransferase activity related genes indicates epigenetic modifications related to abiotic stress and the genes that may act in secondary cell wall biogenesis could lead to strengthening of the cell wall and reduction in pore size in stress conditions. At the 2 d time point, significant biological processes included regulation of transcription (89), embryo development ending in seed dormancy (15), multicellular organism development (13), methylation (9) and response to jasmonic acid (8). At all the three time points, we observed an enrichment of genes encoding TFs which indicates that these are key regulators of a set of genes involved in transcriptional reprogramming during cold acclimation. Concerning the category cellular components, we observed the highest number of targets associated with the nucleus (136 target genes) which is in line with the categories outlined before and underlines the massive processes of transcriptional regulation in response to cold acclimation (Fig. 6).

### Construction of a gene regulatory network (GRN)

To understand the possible interactions and contributions of the major gene regulatory classes, we reconstructed a miRNA and TF regulatory network (Additional file 7, Data S1). The network comprises direct miRNA-mediated target control, miRNAs that regulate transcripts encoding TFs regulating their downstream targets (indirect targets), and TFs which are not miRNA-controlled but regulating miRNA regulated downstream targets (direct targets). To construct the final network, we considered the generated miRNA and mRNA expression data and analyzed all miRNA targets that were predicted using the psRNATarget tool together with publicly available information of TF binding sites (TFBS) and downstream targets. We included experimentally validated regulatory connections from Arabidopsis Transcriptional Regulatory Map [79] and Agris [80]. Further, we included TF target interactions with high confidence from PlantRegMap [81] only considering TFs with different criteria of binding site conservation. First criterion includes TFs and their targets whose binding sites lie within conserved elements of different plant species (CE) whereas the second criterion includes TFs and targets whose binding sites were found to be conserved in different plant species when scanned for conservation of TFBS (FunTFBS).

The validity of the connections in the network was tested by predicting miRNA- and TF-controlled target mRNA expression levels based on miRNA or TF expression levels at a given time point. Here, the prediction power is used as an indicator for the reliability of regulatory links in the network and is calculated by Pearson

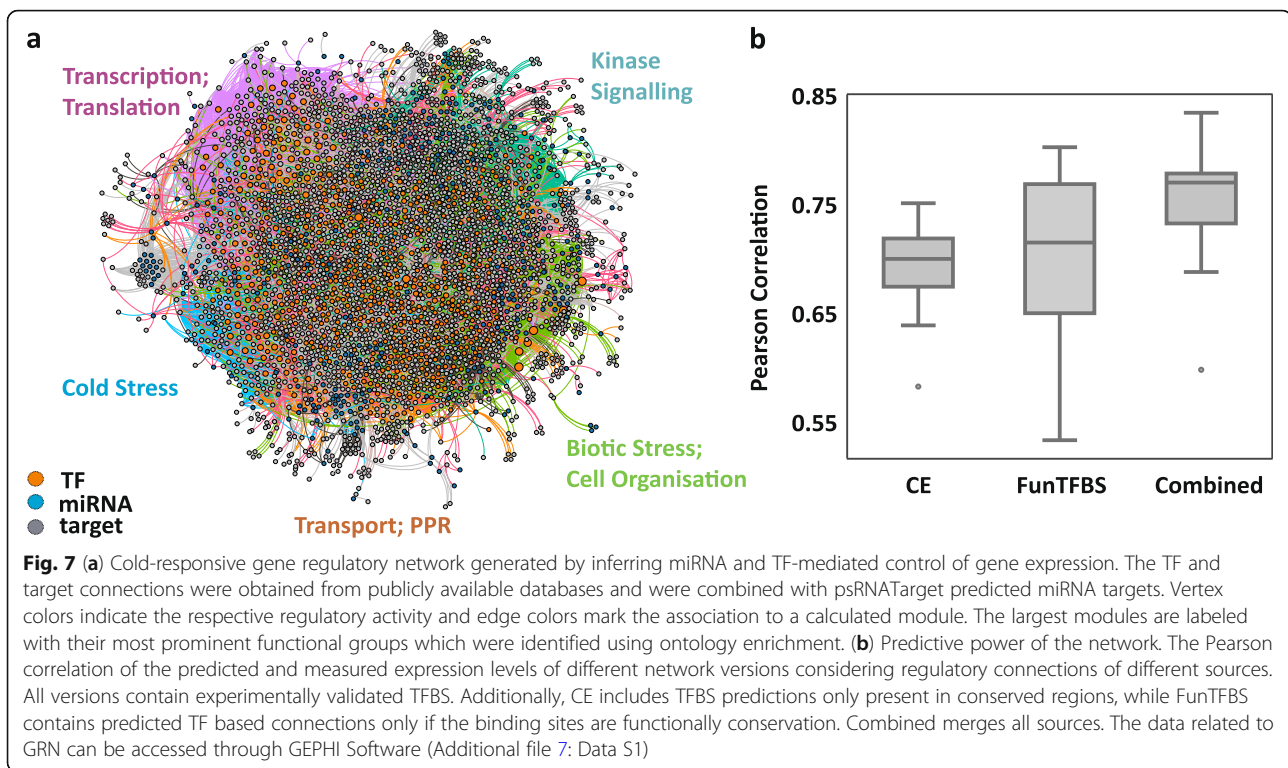
correlation between the predicted and the measured mRNA expression level (Fig. 7 b). We tested the predictive power of the three different network versions to ensure maximal information in the model. Here, the combined version is able to explain on average 77% of the change in target gene expression (0.77 Pearson correlation coefficient) and was considered for further investigation.

This resulting network model contains 350 miRNAs classified into 166 families and consisting a total of 657 TFs belonging to 38 families that either activate or repress 2420 downstream target genes. In total, there are 36,523 regulatory relationships out of which 3846 are miRNA based whereas the remaining 32,677 are TF-based (Fig. 7 a, Additional file 7: Data S1, Additional file 8: Fig. S1).

After validation of the network reconstruction we analyzed the network modularity. Modules are clusters of nodes which are closely connected to each other compared to other nodes in the network. In biological systems, nodes of one module are often co-regulated and closely associated in function. Modules can therefore be interpreted as the functional units of the cell [82]. By using the community detection method [83], we found 17 modules. Functional enrichment using GO and MapMan ontology revealed signaling, transport, cold and biotic stress components, RNA and protein synthesis and cellular organization to be overrepresented in five major network modules.

A cold responsive subnetwork (Fig. 8 a, Additional file 7: Data S2, Additional file 9: Fig. S2) comprising targets of differentially expressed miRNA and targets encoding TFs and their downstream targets was extracted from the GRN. The depicted targets are differentially expressed in at least one of the time points and the extracted network is comprised of 830 nodes and 1332 edges. We observed 103 mature miRNAs and 58 TFs to be involved in the regulation of 669 direct and indirect targets. The functional enrichment revealed a predominant regulation of genes related to cold acclimation, transcription/translation, biotic stress/cell organization, signaling/protein degradation and cell wall/lignin synthesis.

We selected two subnetworks, for miR319 which was DE at all the three time points and miR858 found to be DE at 6 h and 2 d. The miRNA-TF subnetwork of these two miRNAs was extracted from the whole network (Fig. 8 b, c) and the depicted targets in the network are DE in at least one of the analyzed time points (Additional file 10: Fig. S3 and Additional file 11: Fig. S4). The miR858 subnetwork consists of 30 nodes and 51 edges. Among its targets miR858 controls the expression of *Tryptophan synthase (TSB1, AT5G54810)* catalyzing tryptophan synthesis that is the precursor of the auxin indole-3-acetic acid [84]. MiR858 also controls a transcript encoding the TF MYB111 (AT5G49330) which modulates the salt stress response by regulating



flavonoid biosynthesis [85] and the heat shock factor *HSFA4A* (AT4G18880) involved in the response to heat stress. We found 25 nodes and 43 edges to be linked with miR319 that mediates regulation of transcripts such as *TRANSPARENT1 TESTA 8* (TT8, AT4G09820) encoding a TF regulating anthocyanin biosynthesis by the control of *dihydroflavonol 4-reductase* [86]. MiR319 also regulates mRNA for the thermotolerance related heat shock factor *HSFB-2b* (AT4G11660) and a transcript coding for Probable pectinesterase/pectinesterase inhibitor 25 (PME25, AT3G10720) that could facilitate cell wall modifications in cold stress.

#### Differentially expressed sRNAs derived from various other RNA classes

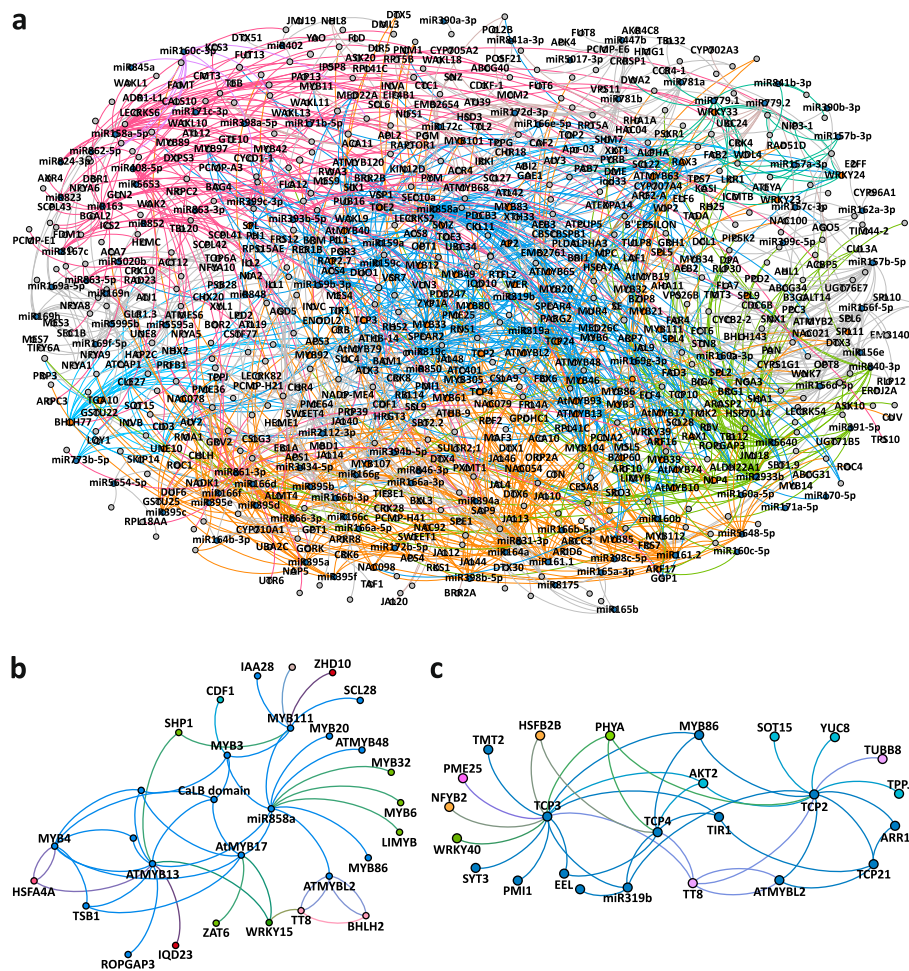
We used our sRNA sequencing data not only to analyze miRNA regulation in response to cold, but also to identify sRNAs derived from other RNA classes which could provide links to their role in cold acclimation. We mapped sRNA reads to publicly available reference databases of lncRNAs, *trans*- and *cis*-NATs pairs, *TAS* and *PHAS* [57, 87–89] and we were able to associate a high number of DE sRNAs to these RNA classes.

#### sRNAs derived from non-overlapping lncRNAs

Here we define non-overlapping lncRNAs as transcripts with a size larger than 200 nt that are single stranded RNA and do not overlap with protein coding transcripts or other non-coding transcripts. In our sRNA data, we

observed 15 non-redundant non-overlapping lncRNA loci that produce DE sRNAs and 13 of these upregulated sRNA production whereas the remaining two downregulated sRNAs in response to cold (Additional file 12: Table S13). However, even if these lncRNAs generate DE sRNAs, the transcript levels of the lncRNAs remained unchanged across the analyzed samples. We found one lncRNA at 3 h, another lncRNA at 6 h and 7 lncRNAs at the 2 d time point of cold treatment that produced DE sRNAs. In addition, we found two lncRNAs differentially producing sRNAs at 3 h as well as 6 h out of which AT5G07745 reduced sRNA production and the other (AT5G04445) upregulated sRNAs at both time points. At 6 h and 2 d we detected four common lncRNAs producing sRNAs with elevated expression levels. The lncRNA AT5G05455 was the only one that produced reduced amounts of sRNAs at the 2 d time point whereas others were upregulated. Single stranded transcripts have the capability to produce fold back structures forming dsRNA which can be processed into small RNAs, but we observed sRNAs produced from sense as well as antisense strands of these lncRNA transcripts. Since these lncRNAs do not overlap with any other transcript and do not have any pairing partners in other genomic loci, it probably indicates that RNA dependent RNA polymerases are involved in the formation of dsRNA from these lncRNA in a primer independent manner that are later converted to sRNAs [9, 90].





**Fig. 8** The extracted cold responsive GRN comprising of direct and indirect targets of DE miRNAs. **(a)** Network of miRNAs and targets that are differentially expressed at any one of the analyzed time points ( $FC \geq 2$  &  $\leq -2$ , Benjamini-Hochberg corrected  $p$ -value  $\leq 0.05$ ). Functional modules associated with cold acclimation, kinase signaling, transcription; translation and transport are represented by blue, dark green, pink, and orange color, respectively. (blue nodes = miRNAs, orange nodes = TFs, gray nodes = targets) **(b)** Subnetwork of miR858. **(c)** Subnetwork of miR319. In **(b)** and **(c)** direct and indirect targets of miRNAs are differentially expressed in at least one of the analyzed time points ( $FC \geq 2$  &  $\leq -2$ , Benjamini-Hochberg corrected  $p$ -value  $\leq 0.05$ ). Curved edges indicate regulatory connections of a regulator and its target. The node colors depict the inferred function based on GO enrichment analyses. Green: biotic stress, cell organization; blue: cold stress; pink: transcription, translation; orange: transport, PPR; dark blue: cell wall, lignin synthesis; red: signaling, protein degradation. The data related to GRN can be accessed through GEPHI Software (Additional file 7: Data S2)

### sRNAs derived from NATs

NATs are pairs of transcripts either non-coding (nc) or protein coding (pc) genes that overlap and form dsRNAs due to sequence complementarity. The pairing of transcripts is possible between nc-nc, nc-pc and pc-pc transcripts and the resulting paired transcript serve as targets for DCL-mediated processing into sRNAs. We found the majority of *cis*- and *trans*-NAT pairs to be produced from pc:pc or pc:nc transcript pairs. In case of pc:nc, the nc pairing partner mostly represents pre-tRNAs or transcripts from TE which also have the capacity to produce sRNAs individually [91–93]. It is known that pre-tRNA and TE-derived sRNAs have the capacity to regulate other transcripts by sequence complementarity which could

indicate their contribution in regulation of cold acclimation related network [92, 94]. Our data set revealed that transcript pairs producing elevated levels of sRNAs in response to cold can have different expression patterns. They can show anticorrelation (one transcript upregulated and the other downregulated), a same tendency of expression (both transcripts either upregulated or down regulated) or no correlation (one transcript regulated and the other remains unchanged). During stress conditions, reverse sequence complementary transcripts of a stress-induced gene and a constitutively expressed gene pair to each other and produce 24 nt and 21 nt siRNAs. The siRNAs produced have the capability to cleave the constitutively expressed transcript resulting in its downregulation

to facilitate stress acclimation. This mechanism represents the classical expression pattern of NATs [54]. The pair is characterized by induced differential expression of nat-siRNAs and anticorrelated expression pattern of the sense and antisense transcripts. We observed abundant sRNAs that were regulated, but their transcript levels remained unchanged. The second most abundant case was an upregulation or downregulation of one of the transcripts whereas the other transcript remained unchanged (Table 2).

Most of the *trans*-NATs gene pairs produced large amounts of sRNAs after 2 d of cold treatment and showed a decrease or no change in the gene transcript levels deduced from the mRNA data. This indicates the possible pairing of both transcripts which are further processed into nat-siRNAs and higher production of these nat-siRNA in the cold acclimation could be required to keep at least one of transcripts at steady levels.

#### *cis*-nat-siRNAs

We found 5, 20 and 100 *cis*-NATs loci (104 non-redundant pairs) at 3 h, 6 h and 2 d, respectively, that produced DE sRNAs from two overlapping transcripts one of which is up- or downregulated whereas the other one remains unchanged (Table 2) (Additional file 12: Table S14). In addition, we detected 24, 34 and 278 *cis*-NATs (308 non-redundant pairs) at 3 h, 6 h and 2 d time point that produced DE sRNAs, but where the cognate overlapping transcripts remained unchanged or could not be detected. Prevalently, we observed that most of the pairs producing *cis*-nat-siRNA were pc:pc transcript

pairs. We found one NATs pair at 6 h and 9 pairs at 2 d resembling the classical mechanism of antisense transcript regulation by nat-siRNAs (Table 3) [54]. We detected a gene pair that gives rise to an increased production of nat-siRNAs and comprises a cold-induced transcript coding for a RAS-Related GTP-Binding Nuclear Protein (*RAN2*, AT5G20020) and a concomitant downregulation of its pairing transcript encoding a Plant Tudor-like RNA-binding protein (AT5G20030). Until now, functional studies on the Plant Tudor-like RNA binding protein are lacking, but *RAN2* is known to be necessary for nuclear translocation of proteins and for RNA export [95]. Another transcript of a salt stress responsive gene encoding an Oleosin-B3-like protein (AT1G13930) [96] which is known to be ABA-induced [97] was also induced by cold in our data and its transcript is able to pair with the transcript of a T-box TF (AT1G13940) to induce production of *cis* nat-siRNAs. Apart from the above mentioned expression patterns of transcripts that differentially regulate siRNA production, we found sRNA producing loci showing same tendency of transcript expression denoted by the upregulation of both pairing transcripts (16 non-redundant pairs; 1, 1 and 15 at 3 h, 6 h and 2 d, respectively) leading to induced sRNA biogenesis. In this category we observed enrichment of pc:pc as well as pc:nc transcript pairs. Prominent examples from our results include the stress-induced pc:pc transcripts *RARE-COLD-INDUCIBLE 2A* (AT3G05880) and *anaphase-promoting complex/cyclosome 11* (AT3G05870) which cause increased *cis*-nat-siRNA production. We also found a cold-induced pc:nc transcript pair coding for a chloroplast beta amylase and a lncRNA, and this upregulated *cis*-nat-siRNAs production consistently at all the three time points. The beta amylase promotes starch degradation into sugars which may act as osmolytes to maintain osmotic balance under cold stress conditions [98].

#### *trans*-nat-siRNAs

We found 38 non-redundant *trans*-NAT pairs (5, 14 and 26 at 3 h, 6 h and 2 d, respectively) that generated DE *trans*-nat-siRNAs from each transcripts. The transcript levels of these 38 gene pairs showed that one of the pairing transcript was either upregulated (5 transcript pairs) or both were unchanged (33 transcript pairs). Out of these 38, we detected four *trans*-NATs gene pairs that generated DE *trans*-nat-siRNA and were common after 3 h (both gene transcripts unchanged) as well as after 6 h (one transcript upregulated and the other one unchanged). We observed 41, 39 and 88 (95 non-redundant) *trans*-NATs gene pairs at 3 h, 6 h and 2 d, respectively, that gave rise to DE *trans*-nat-siRNAs from the overlapping region of two transcripts having unchanged or undetected transcript levels (Table 2). We

**Table 2** Overview of DE nat-siRNAs including expression analysis of the underlying *cis*- or *trans*-transcript pairs

Time points	sRNA clusters	↑↑	↑—	↓—	—	↑↓
3 h <i>cis</i> -NATs	up	1	5		3	
	down				21	
6 h <i>cis</i> -NATs	up	1	20		26	1
	down				8	
2 d <i>cis</i> -NATs	up	15	93	7	270	9
	down				8	
Non-redundant		16	104		308	9
3 h <i>trans</i> -NATs	up				8	
	down				33	
6 h <i>trans</i> -NATs	up		6		24	
	down	2	7		15	
2 d <i>trans</i> -NATs	up		9	11	75	
	down		1	13	13	
Non-redundant		2	14	18	95	

The two symbols in the five columns at the right represent the pairing *cis*- or *trans*-transcript partners and indicate their expression as follows: ↑ = upregulated, ↓ = downregulated, — = unchanged (unchanged refers to FDR > 0.05 and/or fold change FC ≥ 2 & ≤ -2).

**Table 3** Examples of cold acclimation induced *cis*-NATs pairs that produce siRNAs resembling the classical *nat*-siRNA expression

Gene 1	sense transcript	FC	Gene 2	antisense transcript	FC
<b>6 h</b>					
AT2G22080	Transmembrane protein	4.75	AT2G22090	UBP1-associated proteins 1A	-1.41
<b>2 d</b>					
AT5G20020	RAS-related GTP-binding nuclear protein 2	2.58	AT5G20030	Plant Tudor-like RNA-binding protein	-2.17
AT3G11830	TCP-1/cpn60 chaperonin family protein	2.86	AT3G11840	Plant U-box 24	-2.9
AT1G03090	Methylcrotonyl-CoA carboxylase alpha chain, mitochondrial / 3-methylcrotonyl-CoA carboxylase 1 (MCCA)	-2.5	AT1G03100	Pentatricopeptide repeat (PPR) superfamily protein	2.47
AT1G72030	Acyl-CoA N-acyltransferases (NAT) superfamily protein	-2.3	AT1G72040	Deoxyribonucleoside kinase	2.44
AT2G40420	Transmembrane amino acid transporter family protein	-2.3	AT2G40430	SMALL ORGAN 4	2.32
AT5G52440	HIGH CHLOROPHYLL FLUORESCENCE 106	-1.7	AT5G52450	MATE efflux family protein	2.9
AT3G16800	E GROWTH-REGULATING 3	-1.4	AT3G16810	Pumilio 24	5.5
AT2G22080	Transmembrane protein	3.13	AT2G22090	UBP1-associated proteins 1A	-1.51
AT1G13930	Oleosin-B3-like protein	4.53	AT1G13940	T-box transcription factor, putative (DUF863)	-1.31

Sense transcript and antisense transcript fold change  $\geq 2$  or  $\leq -2$ , Benjamini-Hochberg corrected *p*-value  $\leq 0.05$  and siRNA expression fold change  $\geq 2$ , Benjamini-Hochberg corrected *p*-value  $\leq 0.05$ .

observed 2, 5 and 23 *trans*-NAT pairs comprising overlapping pc:pc transcript that generate DE *trans*-nat-siRNAs. We found one pc:pc NAT pair that produced reduced nat-siRNAs at 3 h, but increased nat-siRNAs at 6 h and 2 d time points. Both transcripts encode ZED related kinases (ZRK 1, AT3G57710 and ZRK 7, AT3G57770) that are known to be induced at high temperature and to inhibit the immune response in the absence of plant pathogens [99]. In our data, the transcript levels of these two genes were unchanged, but the generation of *trans*-nat-siRNAs from the two overlapping transcripts might be important to keep the transcripts at a steady-state level. After 2 d of cold treatment, we found a pc:pc *trans*-NAT pair that led to increased *trans*-nat-siRNA production from transcripts encoding Plastid Redox Insensitive (PRIN2, AT1G10522) and prolamins like protein (AT5G53905), but the transcript levels for these two genes remained unchanged. It is known that PRIN2 is a plastid protein involved in redox-mediated retrograde signaling and is required for light-activated PEP-dependent transcription. Another similar example comprises a ncRNA (AT1G70185) and a transcript for a hypothetical protein (AT5G53740) that produce high amounts of *trans*-nat-siRNAs, but their transcript levels were unchanged. Apart from pc:pc pairs, we detected pc transcripts that are able to pair with distinct pre-tRNA. In particular, 7 pc transcripts pairing with 36 pre-tRNA transcripts produced DE

*trans*-nat-siRNAs at 3 h, 10 pc transcripts paired with 46 pre-tRNAs at 6 h and 15 pc paired with 82 pre-tRNAs after 2 d of cold treatment. The majority of the *trans*-NAT gene pairs comprised a nc transcript partner encoding a pre-tRNA or RNA deriving from TE. We found a large number of pc:nc pairs that generated DE sRNAs (41, 37 and 65 loci at 3 h, 6 h and 2 d, respectively) where the transcripts levels were undetected or unchanged. There is a possibility that these pc:nc NATs pairs produce sRNA from the double stranded region of two completely or partially overlapping transcripts, which can be referred as *trans*-nat-siRNAs or these could be derived from single stranded region of two partially overlapping tRNA or TE transcripts (Additional file 12: Table S15). In particular, we observed 1, 8 and 17 pc:nc *trans*-NATs pairs at 3 h, 6 h and 2 d, respectively, that produced DE sRNAs from TE transcripts. One widely known example for a TE-derived siRNA is siRNA854 which shows partial complementarity to the 3' UTR of its target encoding an RNA-binding protein involved in stress granule formation known as UBPIb transcript [100]. We also detected TE-derived sRNAs that are able to target mRNA transcripts to promote cold treatment adaptation. Concerning the *trans*-nat-siRNA producing loci we found 13 transcript pairs after 6 h and 34 pairs after 2 d time that produced DE *trans*-nat-siRNAs where one of the transcripts from each pair was either up- or downregulated and the pairing



partner remained unchanged. The time profile revealed that the highest number of DE *trans*-nat-siRNAs were identified after 2 d indicating *trans*-nat-siRNA mediated regulation of gene expression seems to be most important for the late response to cold acclimation.

### Pha-siRNA

At 6 h time point we identified upregulated sRNAs that were derived from a transcript coding for a mitochondrial PPR protein (AT1G63070) and this was already shown to produce pha-siRNAs [101] (Additional file 12: Table S16). Despite the increasing abundance of the pha-siRNAs we were not able to detect the respective PPR transcript in the mRNA data. The most abundant sRNAs were 21 nt in size followed by 22 nt sRNAs generated from this PPR transcript. The 21 nt pha-siRNAs are known to be loaded into the argonaute and RNA-induced silencing complex to mediate cleavage of mRNAs targets. We performed a target prediction for the 21 nt pha-siRNA with psRNATarget applying stringent parameters and identified putative target transcripts that encode other PPR and TPR proteins, the photosystem II subunit QA (AT4G21280), RNA processing factor 2 (AT1G62670) and HVA22 Homologue A (AT1G74520). The RNA processing factor 2 also belongs to a class of PPR protein which facilitates RNA processing in mitochondria [102]. The photosystem II subunit QA is a component of the electron transport chain and the HVA22 Homologue A protein with an unknown function was previously shown to be ABA and stress inducible [103]. In agreement with the observed upregulation of the pha-siRNA we found the transcript levels of one of the putative targets encoding a PPR protein (AT1G18485) to be significantly downregulated.

### Discussion

Our study aims to provide insights into the cold-responsive regulation of different classes of sRNAs and their impact on the control of either the transcripts underlying sRNA production or the control of transcripts targeted by the sRNAs. We combined sRNA sequencing together with sequencing of mRNAs and lncRNAs to correlate changes in mRNA/lncRNA steady state levels to changes in sRNA expression. We observed classical cold stress related marker genes to be upregulated in the mRNA sequencing data which were found to be differentially expressed in a previous study (Lee et al. 2005) (Additional file 13: Table S17). Over the time course of cold treatment, we observed an overall reduction of sRNAs produced from RNA classes such as miRNAs, *trans*- and *cis*-NATs-pairs and lncRNAs. To exclude that these changes are not caused by altered levels of the major components involved in sRNA biogenesis we analyzed the levels of transcripts encoding

sRNA biogenesis associated proteins such as Hua Enhancer 1 (HEN1), RNA dependent RNA polymerase (ATRDR1–6), DCL1–4, HST1, HYL1, Serrate and Suppressor of Gene Silencing 3 (SGS3). Their levels remained unaffected during the time course of cold treatment and we speculate that the reduced sRNA production could be due to a reduced transcription of sRNA precursor transcripts in response to cold acclimation.

### Analysis of miRNAs and their putative targets

We analyzed DE miRNAs since these are powerful regulators of gene expression and are involved in the control of nearly all cellular pathways [104]. We found 107 DE miRNAs over the time course of the treatment and compared our results to previously reported cold-responsive miRNAs in *A. thaliana* [32, 35, 71]. Baev et al. (2014) treated plants at 4 °C for 24 h and sequenced the RNA from rosette leaves and detected 44 DE miRNAs. We found an overlap of 7 miRNAs following the same expression pattern and the majority of these were DE after 2 d of cold treatment. Similarly, Liu et al. (2008) subjected plants to 4 °C, isolated RNA from whole plant tissues and detected 11 DE miRNAs through microarray experiments. We detected 5 of these 11 miRNAs following the same expression pattern. Sunkar et al. (2004) studied DE miRNAs from whole plants treated at 0 °C for 24 h and two miRNAs were also identified as DE miRNAs in our study. We found 14 out of 107 DE miRNAs to be previously identified in *A. thaliana* in cold stress and these comparisons show that there is limited overlap between the different studies which might be due to the applied temperature, duration of the treatment or plant tissue types used in the studies. Several miRNAs such as miR167c, miR168, miR397, miR389, miR400, miR837-5p, miR838, and miR857 were reported to be cold stress responsive in other studies, but were not identified to be differentially expressed in this study [32, 35, 71].

We analyzed the psRNATarget tool predicted putative miRNA targets of the DE miRNAs and found 96, 173 and 267 miRNA target pairs at 3 h, 6 h and 2 d time points, respectively, which reflects the importance of miRNAs in regulating the transcriptome at prolonged cold treatment. Typically, the alterations in miRNA expression affect the abundance of target genes via cleavage of the target transcript after complementary pairing. The responses of several abiotic stresses are regulated by common mediators that facilitate cross talk of multiple signaling pathways [105]. To maintain the temporal and spatial expression of stress-related genes, the regulatory factors comprising TFs and sRNAs are extremely essential. Among the predicted targets of the DE miRNAs, we found mRNAs encoding TFs such as NFY, MYB, TCP

and HSFs. The GO enrichment of all predicted miRNA targets showed that the highest number of targets are associated with the nucleus (136 mRNAs) and 85 of these encode TFs. Some miRNAs were not associated with anticorrelated targets, but their expression pattern supports the findings of previous cold-related studies such as miR161.1 and miR159b, which were found to be downregulated at the 6 h time point. Studies with *SNRK1* overexpression lines showed reduced miR161 and miR159b promoter activity and lowered transcript levels of the respective *MIR* precursors that is likely to cause reduced miR161 and miR159b levels [106]. Plants have a multitude of TFs that are necessary for growth and stress responses and we predicted 85 targets of DE miRNA that encode TFs. We predicted TCP2 (AT4G18390) and TCP4 (AT3G15030) to be targeted by miR319 and which is consistent with previous studies in *A. thaliana* and sugarcane [107]. All miR319 isoforms were downregulated after 2 d of cold treatment which is consistent with a study in rice, where miR319 was downregulated and its target *TCP21* was upregulated by cold treatment [108]. We observed a similar downregulation of miR319 and concomitant upregulation of its targets *TCP2* and *TCP4* after 2 d of cold treatment.

MYB TFs are known to facilitate cell proliferation and to control phenylpropanoid metabolism and hormone responses [109]. We observed upregulation of miR858 and a corresponding downregulation of its putative targets *MYB48*, *MYB34* and *MYB20*. Apart from TFs, targets of miRNAs also comprise transcripts for epigenetic regulators such as methyl transferases. miR163 was upregulated after 6 h and downregulated after 2 d of cold treatment. One of its targets coding for a S-adenosyl-L-methionine-dependent methyltransferases superfamily protein (AT1G15125) was downregulated after 6 h and another target encoding a N2, N2-dimethylguanosine tRNA methyltransferase (AT5G15810) was upregulated after 2 d of cold treatment. The tRNA methyltransferase (AT5G15810) was shown to cause stress-related N2, N2-dimethylguanosine ( $m^2_2G$ ) modification in tRNAs of *A. thaliana* [110]. Usually, tRNA nucleotide modifications occur within tRNAs during their maturation and processing and these modifications are biomarkers of specific stresses and were observed to be induced in response to oxidizing agents [111]. It is also known that stress-induced epitranscriptomic changes regulate tRNA stability, translation initiation, and microRNA-based regulation of transcripts [111].

#### miR159 alters mitochondrial protein import and ethylene biosynthesis

Similarly, miR159 isoforms were upregulated at 3 h, but downregulated after 2 d of cold treatment. The putative target transcript of miR159 encoding a mitochondrial

translocase TIM-44 related protein (AT5G27395) was anticorrelated with 1.4 fold downregulation at 3 h and 2.8 fold upregulation after 2 d. Since mitochondrial proteins are translated in the cytosol and require import into the mitochondria, our results suggest miRNA-mediated regulation of TIM-44 that may lead to altered mitochondrial protein import during cold treatment. It is known that environmental stresses inhibit and stimulate protein import [112]. TIM44 recruits mitochondrial HSP70 and facilitates the import of proteins containing a transit peptide from the inner membrane into the mitochondrial matrix [113]. miR159 is also known to target RNAs coding for MYB TFs, an aminocyclopropane-1-carboxylate synthase (ACC synthase) and proteins of the *Small Auxin-Up RNA (SAUR)* family [114]. Consistent with the previous findings, the upregulation of miR159 was accompanied by a downregulation of 13 *SAUR* mRNAs and a transcript for an ACC synthase (AT4G37770) that is required for ethylene biosynthesis which is known to be a negative regulator of freezing tolerance [115]. Thus, miR159-mediated downregulation of *ACC synthase* observed in our study suggests a reduced ethylene biosynthesis and increased transcription of *CBF* genes.

#### miR395c targets an mRNA for a mg chelatase that promotes thermogenesis in cold acclimation

miR395c was found to be downregulated after 6 h of cold treatment and its putative target coding for the Mg chelatase subunit H was concomitantly upregulated. The Mg chelatase is a multifunctional protein involved in chlorophyll synthesis catalyzing the insertion of  $Mg^{2+}$  ions into protoporphyrin IX to produce Mg protoporphyrin IX (Mg-Proto-IX) [116]. A recent study confirmed the role of Mg-Proto-IX-derived signals in inducing the gene Alternative oxidase 1a (*AOX1a*) [117]. *AOX1a* reduces  $O_2$  to  $H_2O$  without pumping protons from the matrix to the inter-membrane space and in turn dissipates excess energy in the form of heat. The generated heat plays a role in thermogenesis during cold stress conditions and promotes stress tolerance. Moreover, the Mg-Proto-IX signals also lead to increased activities of antioxidant enzymes that add to the maintenance of redox equilibrium in cold stress [118].

#### A putative target of miR408 coding for a galactose oxidase/kelch repeat protein could induce acclimation in an ABA-dependent manner

Interestingly, miR408-5p was upregulated at all analyzed time points. A chickpea *MIR408* overexpression line subjected to drought stress showed reduced levels of its target coding for plastocyanin. The lack of plastocyanin caused an accumulation of copper and increased levels of copper were shown to cause



upregulation of drought responsive genes such as DREB factors and induced their downstream genes *COR47/RD17* and Low Temperature-Induced 78/Responsive to desiccation 29A (*LTI78/RD29A*) [119]. Similarly, we observed upregulation of miR408-5p, transcripts of DREBs and their downstream transcripts *COR47/RD17* and *LTI78/RD29A* [120, 121]. Further, *MIR408* overexpression lines showed an increased efficiency of photosystem II, reduced electrolyte leakage and lipid peroxidation and increased chlorophyll fluorescence resulting in enhanced cold tolerance due to reduced ROS levels [122]. We predicted a miR408-5p target coding for a galactose oxidase/kelch repeat superfamily protein (AT1G67480) that was found to be downregulated at 6 h and 2 d time points indicating cleavage of the mRNA transcript. Song et al. (2013) studied miR394 and one of its targets coding for the galactose oxidase kelch family protein LCR (Leaf Curling Responsiveness) in *A. thaliana* *MIR394* overexpression and *lcr* mutant lines. They demonstrated upregulation of miR394 and downregulation of *LCR* in the presence of ABA indicating their regulation in salt and drought stress. Other galactose oxidase kelch family proteins such as ZEITLUPE (AT5G57360) have been observed to be reduced at low temperatures [123] and KISS ME DEADLY (AT1G80440) was downregulated to induce UV tolerance [124]. There is a possibility that the putative target galactose oxidase/kelch repeat superfamily protein (AT1G67480) could also mediate cold tolerance in an ABA-dependent manner by its downregulation through miR408-5p [45].

#### miRNA-mediated inhibition of chlorophyll biosynthesis and flowering in cold

miR171-3p was downregulated at the 2 d time point and its cognate mRNA target encoding the GRAS domain TF Scarecrow-Like 27 (AT2G45160) was upregulated. It is known that *SCL27* binds to the promoter of the *PORC* gene (protochlorophyllide oxidoreductase) through GT *cis*-element repeats and represses its expression causing reduced chlorophyll synthesis [125]. The upregulation of *SCL27* due to reduction in miR171 levels could facilitate the cold treatment imposed inhibition of chlorophyll biosynthesis.

We detected miR156/157 isoforms to be upregulated at the 2 d time point accompanied with downregulation of their target *SPL3* (Squamosa Promoter Binding Protein-Like 3). It has been shown that overexpression of *MIR156a* maintains reduced levels of *SPL3* transcripts which leads to delayed flowering in *A. thaliana* [126]. In contrast, miR172c was downregulated and its putative target encoding RAP2.7 also known as Target of Early Activation Tagged 1 (TOE1) was upregulated. *A. thaliana* *TOE1* overexpression lines also showed delayed

flowering [127] and it is possible that miR156 and miR172c regulate transcript levels of *SPL3* and *TOE1* under cold treatment to inhibit flowering.

#### A cold-responsive gene regulatory network indicates importance of miRNA-TF-mRNA interaction

By combining the temporal miRNA and mRNA expression data with publicly available knowledge about regulatory binding behavior of miRNAs, TFs and their downstream target genes, we were able to construct a cold-related GRN of *A. thaliana*. In the resulting GRN we observed different modes of target regulation with respect to miRNAs and TFs both regulating direct targets and miRNAs that regulate TF transcripts and thus control additional targets in an indirect manner. A large number of connections was observed between miRNAs and their direct targets, but the number of affected targets increased when miRNA-targeted TFs were included into the network. This indicates that TFs act as the central nodes for relaying information from miRNAs to several TF-affected targets. The extracted cold responsive GRN revealed an overrepresentation of distinct functional modules such as cold stress, biotic stress and cell organization, transcription and translation, transport and PPR, cell wall and lignin synthesis, signaling and protein degradation. This indicates that miRNA-regulation seems to be important to control major cellular pathways that are known to be involved in cold adaptation. The complete GRN as well as specific subnetworks can be used to study the regulatory relationships of miRNA, TFs and their direct and indirect targets to explore putative novel interacting regulatory components that facilitate cold acclimation.

#### Differentially expressed sRNAs derived from other RNA classes

We further investigated sRNAs derived from other RNA classes such as lncRNA, *cis*- and *trans*-NATs, *TAS* and *PHAS*. We found 15 non-redundant, non-overlapping lncRNAs that produced DE sRNAs during the course of cold treatment. Since 12 of these lncRNA transcripts were not detected by RNAseq and 3 were not DE, we speculate that the lncRNA transcripts are efficiently processed into sRNAs to repress their transcript levels. Such an autoregulatory mechanism has been shown in rice where the lncRNA *Long day specific male fertility associated RNA* (*LDMAR*) was able to produce *Psi-LDMAR* siRNAs that were able to repress their parent *LDMAR* transcript by RNA-dependent DNA methylation (RdDM) [128].

Besides non-overlapping lncRNAs, we found 429 non redundant *cis*-NATs and 179 non redundant *trans*-NATs pairs producing DE siRNAs with a high proportion of pc:nc and pc:pc transcript pairs. DE sRNAs

derived from *cis*-NATs have been identified in *A. thaliana* subjected to drought, cold and salt stress treatments [87]. Zhang et al. (2012) grew seedlings for 29 days at 23 °C and shifted them to 5 °C for 24 h and we detected three *cis*-NATs pairs that were reported in this study to give rise to cold-induced nat-siRNAs. One transcript pair, AT5G15845 (ncRNA) and AT5G15850 (CONSTANS-like 1) showed the same pattern of nat-siRNA production as reported for cold and salt stress and the transcript levels of both genes as well as the nat-siRNAs were upregulated [87]. Another transcript pair, AT5G19220 (ADP-glucose pyrophosphorylase) and AT5G19221 (ncRNA) showed unchanged transcript levels, but elevated nat-siRNA production. The second pair showed less normalized reads in untreated samples compared to cold, salt and drought stress in Zhang et al. (2012). Another NATs pair comprising AT3G22120 (Cell wall-plasma membrane linker protein homolog) and AT3G22121 (ncRNA) led to increased nat-siRNA production. The same gene pair was found to generate reduced nat-siRNA in the previous study in response to cold, but produced elevated nat-siRNAs under salt stress [87].

We observed a predominance of pc:nc gene pairs with pre-tRNA or TE as the non-coding transcript partner. We found a large number of pre-tRNA transcripts pairing with protein coding transcripts and producing siRNAs from one or both pairing transcripts. Several pre-tRNA transcripts are able to pair with an mRNA encoding a Gly-Asp-Ser-Leu (GDSL)-like Lipase/Acylhydrolase superfamily protein (AT5G55050) and a GDSL type lipase gene in pepper has been shown to be involved in drought tolerance, the expression of ABA-inducible genes and oxidative stress signaling [129]. Transcripts encoding F-Box containing proteins (AT2G33655, AT1G11270, AT2G16365) that are known to be co-expressed with several abiotic stress related genes [130] or to activate stress-responsive genes [131] showed pairing with pre-tRNA transcripts to produce *trans*-nat-siRNAs. With respect to the expression pattern of the pairing transcripts and the resulting nat-siRNA it is possible that the siRNAs are produced from the pre-tRNA alone or they are processed from a dsRNA formed by pairing of pre-tRNA and the protein coding transcript. tRNA-derived small RNAs (tsRNAs) were initially thought to be degradation products of endonucleases, but recent advances suggest their functional role in the maintenance of genome stability, epigenetic inheritance, stress response and cell proliferation [132]. Studies in other organisms suggest that the expression of these sRNAs referred to as transfer RNA-derived fragments (5'tRF and 3'tRF) can be related to the quality control of protein synthesis [133, 134]. Previous experiments in *A.thaliana* and human suggest that the tRNA-derived sRNA biogenesis depends on the miRNA pathway [135] and tRFs target transcripts of TE to promote genome

stability [91, 136]. A recent study confirmed the loading of 19–25 nt tRFs into AGO proteins suggesting a role of tRNA produced sRNAs in post-transcriptional gene silencing [94, 137–140]. German et al. (2017) observed the accumulation of 19 nt tRNA-derived sRNAs from the 5' end of mature tRNA transcripts in *A. thaliana* pollen. It was concluded that tRFs are processed similar to miRNAs since there was a reduction in tRF accumulation in a *ddm1/dcl1* double mutant. tRFs and TE-derived sRNAs have been observed to be DE in barley in the presence and absence of phosphorous [141] and in response to phosphate deficiency in *A.thaliana* [142]. Moreover, recently a new class of DCL-independent siRNAs termed sidRNAs were identified that are incorporated into AGO4 and trigger de novo methylation in *A. thaliana* [143] suggesting similarity to tRFs. Besides tRNAs, we detected differential regulation of *trans*-nat-siRNAs derived from transposons containing Ty3 Gypsy, CACTA and Ty1 Copia elements. TE-derived siRNAs can cause DNA methylation or induce repressive histone tail modifications to repress TE loci [144]. Furthermore, in *A. thaliana* TE-derived siRNAs can also target protein coding genes. For example the TE-derived siRNA854 was found to control *UBP1* transcript level that encodes Upstream Binding Protein 1a component of plant stress granules [100]. We found 4, 6 and 26 hypothetical protein coding transcripts pairing with TE encoded transcripts, pseudogene RNAs, mRNA and non-coding RNA transcripts at 3 h, 6 h and 2 d time point, respectively, indicating an involvement of several RNA classes in the adaptation to cold treatment.

In addition to nat-siRNAs derived from pc:nc pairing transcripts, we also identified pc:pc *cis*- and *trans*-NATs pairs that produced siRNAs and we observed an increasing number of nat-siRNAs over the time course of the treatment. We detected elevated expression of nat-siRNAs from 9 *cis*-NAT pairs in response to cold where the overlapping transcripts underlying nat-siRNA production follow the classical expression pattern of a nat-siRNA regulon [54]. This is characterized by an increased expression of nat-siRNAs in response to a stimulus due to an elevated transcription of one of the pairing partners that causes downregulation of the cognate partner transcript. We observed cold-induced upregulation of one transcript together with the repression of its cognate pairing transcript and these gene pairs comprised the transcripts *RAN2 GTPase* (AT5G20020) and *Plant Tudor-like RNA-binding protein* (AT5G20030), *TCP-1 chaperonin family protein* (AT3G11830) and *plant U-box 24* (AT3G11840) and *PPR* (AT1G03100) pairing with *mitochondrial/3-methylcrotonyl-CoA carboxylase 1* (AT1G03090). The inverse expression pattern of these pairing transcripts was accompanied by the induction of *cis*-nat-siRNAs in cold treatment. An ideal example is represented by the cold

responsive upregulation of an mRNA encoding a MATE efflux protein (AT5G52450) that is involved in xenobiotic detoxification, disease resistance, and the control of phytohormones and its pairing partner *High Chlorophyll Fluorescence 106* (HCF106, AT5G52440) that displays a concomitant downregulation. Until now, functional studies on the putative MATE efflux protein are lacking whereas the overlapping transcript encoding HCF106 protein is well characterized. HCF106 is a chloroplast thylakoid protein and imports proteins into the thylakoid lumen. The *hcf106* knockout mutants are albino mutants and seedling-lethal, whereas weaker T-DNA alleles are paler in color and display reduced stomatal aperture and reduced water loss and hence cause elevated dehydration tolerance [145]. The production of nat-siRNAs from the two transcripts resulting in elevated levels of the MATE transcript and downregulation of HCF106 transcript suggests a cold-responsive regulatory mechanism which could act in cold acclimation.

Based on our results, we conclude that cold treatment leads to considerable changes in sRNA levels that are likely to contribute to changes in gene expression that underlie cold acclimation in *A. thaliana*. The combination of multilevel high throughput sequencing and bioinformatics analysis proved to be a powerful tool to create a regulatory network of sRNAs and mRNAs responsive to cold stress. A high number of miRNAs were DE and their predicted targets include a large number of mRNAs encoding TFs, PPR and TPR proteins that act in the regulation of gene expression and protein biosynthesis, respectively, and transcripts encoding important enzymes that act in cold acclimation. Along with miRNAs, large numbers of sRNAs were produced from lncRNAs and transcripts of *cis*- and *trans*-NATs pairs indicating a strong impact of all sRNA classes in cold adaptation.

## Conclusions

According to this study in *A. thaliana*, miRNAs and sRNAs derived from, *cis*- and *trans*-NAT gene pairs and from lncRNAs play an important role in regulating gene expression in cold acclimation. The gene regulatory network constructed provides substantial information related to the interaction of miRNA and their associated direct and indirect targets. Overall, this study provides a fundamental database to deepen our knowledge and understanding of regulatory networks in cold acclimation.

## Methods

### Plant material and stress treatment

Seeds of *A. thaliana* ecotype Columbia (*Col-0*) were sown at a high density (ca. 50 seeds on 9 × 9 cm pots) with soil substrate and stratified at 4 °C for 2 d in the dark. Following stratification, the pots were transferred to LED-41 HIL2 cabinets (Percival, Perry, USA) and

cultivated under control conditions with a light / dark regime of 16 h light (80 μmol photons m<sup>-2</sup> s<sup>-1</sup>; corresponding to 18% of blue and red channel) at 22 °C followed by 8 dark at 18 °C for 14 d. Plants serving as controls remained under these condition whereas plants subjected to cold treatment were transferred 4 h after the onset of light at continuous 4 °C with diurnal light intensity of 35 μmol photons m<sup>-2</sup> s<sup>-1</sup>. The cold treatment was performed in three independent subsequent experimental replicates using the same growth chamber with identical settings. The aerial tissues from three experimental replicates of cold-treated as well as control samples were harvested after 3 h, 6 h, and 48 h (2 d).

### RNA isolation and sRNA sequencing

The total RNA from the biological triplicates of each sample were isolated using TRI-Reagent (Sigma) according to the manufacturer's instructions. For each mRNA and lncRNA library including polyA-tailed lncRNAs, 10 μg total RNA was vacuum dried with RNA stable (Sigma-Aldrich). The libraries were prepared by Novogene (China) using the Next Ultra RNA Library Prep Kit (NEB). The libraries were strand-specifically sequenced as 150 bp paired-end on a HiSeq-2500 platform with at least 15 million read pairs per library.

For each sRNA library 50 μg of total RNA was separated on a 15% native polyacrylamide gel. The ZR small-RNA Ladder (Zymo Research) served as RNA size marker and sRNAs corresponding to 17–29 nt were excised from the gel. The gel pieces were transferred into a LoBind Eppendorf tube and crushed using a disposable polypropylene pestle. 0.3 M NaCl was added to immerse the gel pieces and the tubes were frozen for 15 min at –80 °C and RNA was subsequently eluted overnight at 4 °C. The buffer was transferred into a Spin-X centrifuge tube filter (COSTAR) and centrifuged for 1 min at 4 °C to remove the gel pieces. RNA was precipitated by adding 2.5 volume of 100% (v/v) ethanol, 1/10 volume of 3 M NaOAc (pH 5) and 1 μl of glycogen (10 mg/ml) and incubation at –80 °C for 4 h. The samples were centrifuged for 30 min with 17,000 × g at 4 °C and the RNAs were washed twice with 80% ethanol, dried at room temperature and resuspended in 7 μl of nuclease free water. RNA concentrations were measured spectrophotometrically and the sRNA fractions were used for library preparation using the NEBNext multiplex small RNA library prep kit Illumina following the manufacturer's protocol with minor modifications. The 3' SR adapter was ligated at 16 °C overnight and the SR reverse transcription primer was hybridized to an excess of 3' SR adapter to prevent adapter dimer formation. After ligation of the 3' SR adapter, the 5' SR adapter was ligated to the RNA and incubated for 1.5 h at 25 °C. PCR amplification of the libraries was performed using



specific index primers for 12 cycles and the cDNA amplicons were separated on a 6% native acrylamide gel at 120 V. The gel was stained with SYBR gold and RNAs with a size between 138 and 150 nt corresponding to adapter-ligated sRNAs with a size between 18 and 30 nt were excised. Gel elution of the DNA was performed as described above except the addition of 1 µl linear acrylamide (5 mg/ml) prior to precipitation to increase the DNA pellet mass. The cDNA library with concentration of at least 8 ng/µl was considered optimum for sequencing. The sRNA libraries were sequenced with an Illumina deep sequencing platform (Illumina HiSeq 1500) with a read length of 50 nt and a minimum of 7 million reads per library.

### Bioinformatic analyses of transcriptomes

The mRNA/lncRNA sequencing data for the triplicates of 3 h, 6 h and 2 d cold-acclimated samples together with the respective controls were analyzed using open web based platform GALAXY (<https://usegalaxy.org/>) [146]. The adapter sequences were trimmed using the FASTQ Trimmomatic tool using the default parameters. To map the raw reads against *A. thaliana* reference genome (<https://www.arabidopsis.org>, release: TAIR10), Tophat tool was used with a maximum intron length parameter of 3000 nt. The Araport11 annotation [147] was used to annotate the transcripts and ncRNA transcripts longer than 200 bp were considered as lncRNAs. We used the FeatureCounts tool to count the number of reads mapped to the reference genome (Additional file 1: Table S1). Using the count file as an input for the DeSeq2 tool of GALAXY, we obtained the final list of genes. All genes were classified based on Araport11 reference annotation (<https://araport.org/>).

The sRNA raw reads were mapped to the TAIR10 (<https://www.arabidopsis.org>, release: TAIR10) reference genome using the Shortstack software [148]. Approximately 80% of the obtained reads efficiently mapped to it (Additional file 1: Table S2). We generated a reference annotation database for sRNAs derived from RNA classes such as miRNA (miRBase version 22.1), lncRNA (Araport11), *trans*- and *cis*-nat-siRNA [57, 87–89], ta-siRNA and phasiRNA [101] that was used to generate read counts of sRNAs obtained from these RNA classes. The counts generated from the triplicates were used for the analysis of differential expression using the DeSeq2 tool in GALAXY and sRNAs having a  $FC \geq 2$  &  $\leq -2$ , Benjamini-Hochberg corrected  $p$ -value  $\leq 0.05$  were considered to be DE. Global comparisons of DE miRNAs were generated using UpSetR package (<https://CRAN.R-project.org/package=UpSetR>).

### cDNA synthesis for stem loop qRT-PCR

cDNA was synthesized using 300 ng of RNA from three biological replicates of treated and untreated samples

[149]. The RNA was treated with DNase I (2 U, NEB) at 37 °C for 30 min to eliminate genomic DNA contamination, the enzyme was heat-inactivated at 65 °C for 10 min and the RNA was reverse transcribed into cDNA by M-MuLV Reverse transcriptase (200 U, NEB) at 42 °C for 30 min. Specific stem loop primers and a universal reverse primer were used for cDNA synthesis (Additional file 14: Table S18). During cDNA synthesis, we added *UBI1* (AT4G36800) specific reverse primer and monitored the successful cDNA synthesis through PCR by using *UBI1* specific gene primers.

### Stem loop qRT-PCR

The Real-time PCR was performed using EvaGreen and sRNA-specific primers (Additional file 14: Table S18). For each sample, the qRT-PCR was performed in three technical replicates and each reaction contained cDNA amounts equivalent to 20 ng/µl of initial RNA. The qRT-PCR program was adjusted to initial denaturation at 95 °C for 2 min followed by 40 cycles of amplification with 95 °C for 12 s, annealing for 30 s and 72 °C for 15 s. The SYBR green signals were measured after each cycle and melting curves were monitored to confirm primer specificities. The  $C_t$  values were used to calculate the expression levels by using  $\Delta\Delta C_t$  method [150]. The expression levels were normalized using *UBI1* housekeeping gene (AT4G36800).

### miRNA target prediction

MiRNA targets were predicted using the psRNATarget prediction tool (2017 Update) [73]. DE miRNAs were used as a query to search against *A. thaliana* protein coding and non-coding transcripts of Araport11 keeping default parameters and allowing calculation of target accessibility (maximum energy to unpair the target site = 25). We used a stringent cut off value 2.5 as the maximum expectation score for selecting our potential targets.

### Gene ontology of miRNA targets

GO analyses were performed with the DAVID Bioinformatics tool [78]. The list of miRNA target genes was provided as an input and the output list contained genes categorized into biological process, cellular compartment and molecular function. We filtered for significant GO terms with Benjamini-Hochberg corrected  $p$ -value  $\leq 0.05$  which was obtained from Fisher's test in all the categories. The dot plot visualizing the GO terms was generated using ggplot2 package (<https://CRAN.R-project.org/package=ggplot2>).

### Construction and validation of the regulatory network model

The gene regulatory network (GRN) was constructed using high confidence experimentally validated regulatory connection from ATRM [79] and Agris [80]. We

did not include all the connections available in PlantReg-Map [151] but the ones which fulfill the criteria of conservation of binding motifs. First criterion includes TF connections whose binding sites lie in the conserved elements of different plant species (motif\_CE) and the second criterion included TF connections whose binding sites were found to be conserved in different plant species when scanned for conservation of TFBSs (FunTFBS) [81]. The TF based regulatory connections following these two criteria were merged with the psRNATarget tool predicted miRNA targets to obtain the full network model. The prediction of target gene expression was performed using the Fast Tree Regression learner from Dotnet.ML version 0.8 [152]. The outcome variable was the FPKM of target gene expressions at the separate time points 3 h, 6 h, and 2 d. As input variables, we used the time point, the expression levels for each regulator familywise aggregated at the respective time and the counts of binding sites of the target gene. Both family assignments for each TF and binding site information for each target were taken from the AtTFDB database [153]. The data related to GRN can be accessed through free visualization Software GEPHI available for download at <https://gephi.org/> (Additional file 7: Data S1, S2).

### Heatmap clustering

The pheatmap function (<https://cran.r-project.org/web/packages/pheatmap/index.html>) of the R package 'Pheatmap' was used to create a heatmap showing hierarchical clustering of differentially expressed miRNAs at the three time points of cold treatment.

### Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s12870-020-02511-3>.

**Additional file 1 Table S1:** Total mRNA sequencing reads mapping to the *A. thaliana* reference genome after adapter trimming in control and cold treated samples (biological triplicates). **Table S2:** Total sRNA sequencing reads mapping to different sRNA producing RNA classes for control and cold treated samples. **Table S3:** sRNA size distribution in reads per million. The size distribution of total sRNAs derived from control and cold treated samples after adapter trimming.

**Additional file 2 Table S4:** Differentially expressed miRNAs during cold acclimation. The three sub-tables depict DE miRNAs at 3 h, 6 h and 2 d, respectively. The miRNAs highlighted in orange belong to evolutionarily conserved miRNA families. **Table S5:** Normalized read counts and fold changes of all miRNAs during cold acclimation. The three sub-tables depict all miRNAs at 3 h, 6 h and 2 d, respectively. **Table S6:** Differentially expressed cold-responsive miRNAs in *A. thaliana*. Fold changes of miRNAs after 3 h, 6 h and 2 d of cold treatment, considered DE when  $\log_2FC \geq 1$  &  $\leq -1$ , Benjamini-Hochberg corrected  $p$ -value  $\leq 0.05$ . Conserved miRNAs are highlighted in bold.

**Additional file 3 Table S7:** List of miRNA-targeted mRNAs predicted using psRNATarget. The sub-tables depict all predicted targets of DE miRNAs at the three time points. A stringent expectation value of 2.5 was used to filter the targets. N/A = No significant fold change. **Table S8:** List of miRNA-targeted ncRNAs predicted using psRNATarget. The sub-tables depict all predicted ncRNA targets of DE miRNAs at the three time points.

A stringent expectation value of 2.5 was used to filter the targets. N/A = No significant fold change.

**Additional file 4 Table S9:** List of all mRNAs generated from mRNA sequencing data. The three sub-tables depict normalized read counts (triplicates) from control and cold treated samples at 3 h, 6 h and 2 d.

**Table S10:** List of all significant DE mRNAs generated from mRNA sequencing data. The sub-tables present all the details from control and cold treated samples after 3 h, 6 h and 2 d.

**Additional file 5 Table S11:** List of 54 targets of differentially expressed miRNAs from all the four subgroups found to be consistently present at all the three time points. The Venn diagram depicts all targets of differentially expressed miRNAs observed after 3 h, 6 h and 2 d.

**Additional file 6 Table S12:** Gene Ontology term enrichment analysis for predicted targets of differentially expressed miRNAs. The sub-tables depict GO terms after 3 h, 6 h and 2 d of cold acclimation.

**Additional file 7.** The data file that can be accessed using free software GEPHI available at <https://gephi.org/> comprising of Data S1: Gene regulatory network in cold acclimation, Data S2: Cold responsive network of the differentially expressed miRNAs.

**Additional file 8 Fig. S1:** Complete gene regulatory network (GRN) of cold acclimation. Overview of the GRN for cold acclimation. All predicted miRNA targets in cold were selected and TFs regulating these targets were inferred. Vertex colors indicate the respective regulatory activity and edge colors mark the association to a calculated module. The biggest modules are labeled with their most prominent functional groups which were identified using ontology enrichment.

**Additional file 9 Fig. S2:** Cold responsive gene regulatory network comprising of direct and indirect targets of DE miRNAs. The miRNAs and the targets are differentially expressed at any one of the analyzed time points ( $FC \geq 2$  &  $\leq -2$ , Benjamini-Hochberg corrected  $p$ -value  $\leq 0.05$ ). Functional modules associated with cold stress; kinase signaling; transcription, translation and transport are represented by blue, dark green, pink, and orange color, respectively.

**Additional file 10 Fig. S3:** Subnetwork of miR858a extracted from the complete network. The direct and the indirect targets of miRNAs are differentially expressed in at least one of the analyzed time points ( $FC \geq 2$  &  $\leq -2$ , Benjamini-Hochberg corrected  $p$ -value  $\leq 0.05$ ).

**Additional file 11** Subnetwork of miR319b extracted from the complete network. The direct and the indirect targets of miRNAs are differentially expressed in at least one of the analyzed time points ( $FC \geq 2$  &  $\leq -2$ , Benjamini-Hochberg corrected  $p$ -value  $\leq 0.05$ ).

**Additional file 12 Table S13:** Differentially expressed sRNAs produced from non-overlapping lncRNAs. The sub-tables depict detailed sRNA and lncRNA transcript sequencing data at 3 h, 6 h and 2 d. **Table S14:** Differentially expressed sRNAs produced from *cis*-NAT pairs. The sub-tables depict detailed sRNA and *cis*-NAT sequencing data at 3 h, 6 h and 2 d.

**Table S15:** Differentially expressed sRNAs produced from *trans*-NAT pairs. The sub-tables depict detailed sRNA and *trans*-NAT sequencing data at 3 h, 6 h and 2 d. **Table S16:** Differentially expressed sRNAs produced from *PHAS* pairs. The sub-tables depict detailed sRNA and *PHAS* transcript sequencing data at 3 h, 6 h and 2 d.

**Additional file 13 Table S17:** List of classical cold responsive genes that were found to be differentially expressed in Lee et al. 2005 and are also differential expression in our study.

**Additional file 14 Table S18:** Sequences of oligonucleotides used in this study to perform stem loop qRT-PCR.

### Abbreviations

sRNAs/ siRNAs: Small RNAs/ small interfering RNAs; miRNA: MicroRNA; ncRNA: Non-coding RNA; lncRNA: Long non-coding RNA; ta-siRNA: Trans-acting siRNA; *cis/trans*-NAT: *Cis/trans*-natural antisense transcript; NAT: Natural antisense transcript; DEG: Differentially expressed gene; TE: Transposable element; TF: Transcription factor; pc: Protein coding; nc: Non-coding; GO: Gene Ontology; OST1: OPEN STOMATA 1; ABA: Absciscic acid; ICE1: Inducer of CBF expression; CBF: C-repeat binding factors; DREB: Dehydration responsive element binding factors; CRT/DRE: Cold response sensitive transcription factors/dehydration responsive elements;

COR: Cold-responsive; ABFs: ABRE-binding factors; RdDM: RNA-directed DNA methylation; FLC: Flowering locus C; DCL1: DICER-LIKE1; LCR: LEAF CURLING RESPONSIVENESS; P5CDH: Delta-pyrroline-5-carboxylate dehydrogenase; SRO5: Similar to Radicle Induced Cell Death One 5; DDM1: Decreased DNA methylation 1; ea-siRNA: Epigenetically activated siRNA; TCP: Teosinte Branched 1, Cycloidea and Pcf Transcription Factor 2; TIM: Translocase Inner Membrane Subunit; RAN2: RAS-Related GTP-Binding Nuclear Protein; PPR: Pentatricopeptide repeat superfamily protein; NFY: Nuclear Factor-Y; HSF: Heat shock factors; ACC synthase: Amino-cyclopropane-1-carboxylate synthase; MATE: Multi-antimicrobial extrusion protein; tRF: tRNA-derived RNA fragments

## Acknowledgements

We thank Martin Simon for technical support and advice regarding sRNA library preparations and Oguz Top for helpful comments on the manuscript.

## Authors' contributions

WF and MAA designed the research; BT performed the research with the help of MAA and KH; BT, MAA, KH and WF analyzed the data; AGM and TK provided the 3 h and 2 d mRNA/lncRNA mRNA sequencing raw data; miRNA-TF network was constructed by HLW and TM; and BT, MAA, HLW and WF wrote the paper. The authors read and approved the final manuscript.

## Funding

The funding for the research conducted was provided by the German Research Foundation (SFB-TRR 175, grants to W.F. project C03, T.M. project D02, and T.K. project C01). The Funding body was not involved in the design of the study and analysis or interpretation of the data and in writing of the manuscript.

## Availability of data and materials

The raw Illumina sRNA and mRNA sequencing data is deposited in NCBI SRA database with the ID PRJNA592037. All raw data used for the analyses in this study is available for reviewers at <https://dataview.ncbi.nlm.nih.gov/object/PRJNA592037?reviewer=1hkljqn6c6qp67vp6p70ra9l59>.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup>Department of Biology I, Plant Molecular Cell Biology, Ludwig-Maximilians-Universität München, LMU Biocenter, Großhaderner Str. 2-4, 82152 Planegg-Martinsried, Germany. <sup>2</sup>Computational Systems Biology, Technische Universität Kaiserslautern, Paul-Ehrlich-Straße 23, 67663 Kaiserslautern, Germany. <sup>3</sup>Department of Biology I, Plant Molecular Cell Biology, Ludwig-Maximilians-Universität München, LMU Biocenter, Großhaderner Str. 2-4, 82152 Planegg-Martinsried, Germany.

Received: 26 February 2020 Accepted: 22 June 2020

Published online: 29 June 2020

## References

- Gornall J, Betts R, Burke E, Clark R, Camp J, Willett K, et al. Implications of climate change for agricultural productivity in the early twenty-first century. *Philos Trans R Soc Lond B Biol Sci*. 2010;365(1554):2973–89.
- Zhu JK. Abiotic stress signaling and responses in plants. *Cell*. 2016;167(2):313–24.
- Solanke AU, Sharma AK. Signal transduction during cold stress in plants. *Physiol Mol Biol Plants*. 2008;14(1–2):69–79.
- Thomashow MF. Role of cold-responsive genes in plant freezing tolerance. *Plant Physiol*. 1998;118(1):1–8.
- Abla M, Sun H, Li Z, Wei C, Gao F, Zhou Y, et al. Identification of miRNAs and their response to cold stress in *Astragalus Membranaceus*. *Biomolecules*. 2019;9(5):182.
- Ding Y. OST1 kinase modulates freezing tolerance by enhancing ICE1 stability in *Arabidopsis*. *Dev Cell*. 2015;32(3):278–89.
- Chinnusamy V, Zhu JK, Sunkar R. Gene regulation during cold stress acclimation in plants. *Methods Mol Biol*. 2010;639:39–55.
- Cuevas-Velazquez CL, Rendon-Luna DF, Covarrubias AA. Dissecting the cryoprotection mechanisms for dehydrins. *Front Plant Sci*. 2014;5:583.
- Devert A, Fabre N, Floris M, Canard B, Robaglia C, Crete P. Primer-dependent and primer-independent initiation of double stranded RNA synthesis by purified *Arabidopsis* RNA-dependent RNA polymerases RDR2 and RDR6. *PLoS One*. 2015;10(3):e0120100.
- Wang DZ, Jin YN, Ding XH, Wang WJ, Zhai SS, Bai LP, et al. Gene regulation and signal transduction in the ICE-CBF-COR signaling pathway during cold stress in plants. *Biochemistry*. 2017;82(10):1103–17.
- Lee SJ, Kang JY, Park HJ, Kim MD, Bae MS, Choi HI, et al. DREB2C interacts with ABF2, a bZIP protein regulating abscisic acid-responsive gene expression, and its overexpression affects abscisic acid sensitivity. *Plant Physiol*. 2010;153(2):716–27.
- Lamke J, Baurle I. Epigenetic and chromatin-based mechanisms in environmental stress adaptation and stress memory in plants. *Genome Biol*. 2017;18(1):124.
- Banerjee A, Wani SH, Roychoudhury A. Epigenetic control of plant cold responses. *Front Plant Sci*. 2017;8:1643.
- Palusa SG, Ali GS, Reddy AS. Alternative splicing of pre-mRNAs of *Arabidopsis* serine/arginine-rich proteins: regulation by hormones and stresses. *Plant J Cell Mol Biol*. 2007;49(6):1091–107.
- Chekanova JA. Long non-coding RNAs and their functions in plants. *Curr Opin Plant Biol*. 2015;27:207–16.
- Li S, Castillo-Gonzalez C, Yu B, Zhang X. The functions of plant small RNAs in development and in stress responses. *Plant J Cell Mol Biol*. 2017;90(4):654–70.
- Ku YS, Wong JW, Mui Z, Liu X, Hui JH, Chan TF, et al. Small RNAs in plant responses to abiotic stresses: regulatory roles and study methods. *Int J Mol Sci*. 2015;16(10):24532–54.
- Ransohoff JD, Wei Y, Khavari PA. The functions and unique features of long intergenic non-coding RNA. *Nat Rev Mol Cell Biol*. 2018;19(3):143–57.
- Wang XQ, Crutchley JL, Dostie J. Shaping the genome with non-coding RNAs. *Curr Genomics*. 2011;12(5):307–21.
- Di C, Yuan J, Wu Y, Li J, Lin H, Hu L, et al. Characterization of stress-responsive lncRNAs in *Arabidopsis thaliana* by integrating expression, epigenetic and structural features. *Plant J Cell Mol Biol*. 2014;80(5):848–61.
- Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. *Mol Cell*. 2011;43(6):904–14.
- Franco-Zorrilla JM, Valli A, Todesco M, Mateos I, Puga MI, Rubio-Somoza I, et al. Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat Genet*. 2007;39(8):1033–7.
- Swiezewski S, Liu F, Magusin A, Dean C. Cold-induced silencing by long antisense transcripts of an *Arabidopsis* Polycomb target. *Nature*. 2009;462(7274):799–802.
- Csorba T, Questa JI, Sun Q, Dean C. Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalization. *Proc Natl Acad Sci*. 2014;111(45):16160–5.
- Matzke MA, Mosher RA. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat Rev Genet*. 2014;15(6):394–408.
- Phil Chi Khang Au ESD. Analysis of Argonaute 4-associated long non-coding RNA in *Arabidopsis thaliana* sheds novel insights into gene regulation through RNA-directed DNA methylation. *Genes*. 2017;8(8):198.
- Mallory AC, Vaucheret H. Functions of microRNAs and related small RNAs in plants. *Nat Genet*. 2006;38(Suppl):S31–6.
- Gusta LV, Trischuk R, Weiser CJ. Plant cold acclimation: the role of abscisic acid. *J Plant Growth Regul*. 2005;24(4):308–18.
- Chen X. MicroRNA biogenesis and function in plants. *FEBS Lett*. 2005;579(26):5923–31.
- de Lima JC, Loss-Morais G, Margis R. MicroRNAs play critical roles during plant development and in response to abiotic stresses. *Genet Mol Biol*. 2012;35(4):1069–77.
- Song G, Zhang R, Zhang S, Li Y, Gao J, Han X, et al. Response of microRNAs to cold treatment in the young spikes of common wheat. *BMC Genomics*. 2017;18(1):212.
- Baev V, Milev I, Naydenov M, Vachev T, Apostolova E, Mehterov N, et al. Insight into small RNA abundance and expression in high- and low-temperature stress response using deep sequencing in *Arabidopsis*. *Plant Physiol Biochem*. 2014;84:105–14.

33. Lee BH, Henderson DA, Zhu JK. The *Arabidopsis* cold-responsive transcriptome and its regulation by ICE1. *Plant Cell*. 2005;17(11):3155–75.
34. Lee H, Xiong L, Ishitani M, Stevenson B, Zhu JK. Cold-regulated gene expression and freezing tolerance in an *Arabidopsis thaliana* mutant. *Plant J Cell Mol Biol*. 1999;17(3):301–8.
35. Liu HH, Tian X, Li YJ, Wu CA, Zheng CC. Microarray-based analysis of stress-regulated microRNAs in *Arabidopsis thaliana*. *Rna*. 2008;14(5):836–43.
36. Mahale B, Fakrudin B, Ghosh S, Krishnaraj PU. LNA mediated in situ hybridization of miR171 and miR397a in leaf and ambient root tissues revealed expressional homogeneity in response to shoot heat shock in *Arabidopsis thaliana*. *J Plant Biochem Biotechnol*. 2013;23(1):93–103.
37. Zhang Y, Zhu X, Chen X, Song C, Zou Z, Wang Y, et al. Identification and characterization of cold-responsive microRNAs in tea plant (*Camellia sinensis*) and their targets using high-throughput sequencing and degradome analysis. *BMC Plant Biol*. 2014;14:271.
38. De Rienzo F, Gabdoulline RR, Menziani MC, Wade RC. Blue copper proteins: a comparative analysis of their molecular interaction properties. *Protein Sci*. 2000;9(8):1439–54.
39. Pilon SEA-GaM. MicroRNA-mediated systemic Down-regulation of copper protein expression in response to low copper availability in *Arabidopsis*. *J Biol Chem*. 2008;283(23):15932–45.
40. Pourcel L, Routaboul J-M, Cheynier V, Lepiniec L, Debeaujon I. Flavonoid oxidation in plants: from biochemical properties to physiological functions. *Trends Plant Sci*. 2007;12(1):29–36.
41. Liang M, Haroldsen V, Cai X, Wu Y. Expression of a putative laccase gene, ZmLAC1, in maize primary roots under stress. *Plant Cell Environ*. 2006;29(5):746–53.
42. Song JB, Huang SQ, Dalmay T, Yang ZM. Regulation of LEAF morphology by microRNA394 and its target LEAF CURLING RESPONSIVENESS. *Plant Cell Physiol*. 2012;53(7):1283–94.
43. Knauer S, Holt AL, Rubio-Somoza I, Tucker EJ, Hinze A, Pisch M, et al. A protodermal miR394 signal defines a region of stem cell competence in the *Arabidopsis* shoot meristem. *Dev Cell*. 2013;24(2):125–32.
44. Song JB, Gao S, Wang Y, Li BW, Zhang YL, Yang ZM. miR394 and its target gene LCR are involved in cold stress response in *Arabidopsis*. *Plant Gene*. 2016;5:56–64.
45. Song JB, Gao S, Sun D, Li H, Shu XX, Yang ZM. miR394 and LCR are involved in *Arabidopsis* salt and drought stress responses in an abscisic acid-dependent manner. *BMC Plant Biol*. 2013;13:210.
46. Dong C-H, Pei H. Over-expression of miR397 improves plant tolerance to cold stress in *Arabidopsis thaliana*. *J Plant Biol*. 2014;57(4):209–17.
47. Peragine A, Yoshikawa M, Wu G, Albrecht HL, Poethig RS. SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in *Arabidopsis*. *Genes Dev*. 2004;18(19):2368–79.
48. Vazquez F, Vaucheret H, Rajagopalan R, Lepers C, Gascoli V, Mallory AC, et al. Endogenous trans-acting siRNAs regulate the accumulation of *Arabidopsis* mRNAs. *Mol Cell*. 2004;16(1):69–79.
49. Guan C, Wu B, Yu T, Wang Q, Krogan NT, Liu X, et al. Spatial auxin signaling controls leaf flattening in *Arabidopsis*. *Curr Biol*. 2017;27(19):2940–50.
50. Moldovan D, Spriggs A, Yang J, Pogson BJ, Dennis ES, Wilson IW. Hypoxia-responsive microRNAs and trans-acting small interfering RNAs in *Arabidopsis*. *J Exp Bot*. 2010;61(1):165–77.
51. Kohei K. TAS1 trans-acting siRNA targets are differentially regulated at low temperature, and TAS1 trans-acting siRNA mediates temperature-controlled At1g51670 expression. *Biosci Biotechnol Biochem*. 2010;74(7):1435–40.
52. Hsieh L-C. Uncovering small RNA-mediated responses to phosphate deficiency in *Arabidopsis* by deep sequencing; 2009.
53. Kumar M, Carmichael GG. Antisense RNA: function and fate of duplex RNA in cells of higher eukaryotes. *Microbiol Mol Biol Rev*. 1998;62(4):1415–34.
54. Borsani O, Zhu J, Verslues PE, Sunkar R, Zhu JK. Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in *Arabidopsis*. *Cell*. 2005;123(7):1279–91.
55. Wight M, Werner A. The functions of natural antisense transcripts. *Essays Biochem*. 2013;54:91–101.
56. Wang XJ, Gaasterland T, Chua NH. Genome-wide prediction and identification of cis-natural antisense transcripts in *Arabidopsis thaliana*. *Genome Biol*. 2005;6(4):R30.
57. Yuan C, Wang J, Harrison AP, Meng X, Chen D, Chen M. Genome-wide view of natural antisense transcripts in *Arabidopsis thaliana*. *DNA Res*. 2015;22(3):233–43.
58. Zhang X, Lii Y, Wu Z, Polishko A, Zhang H, Chinnusamy V, et al. Mechanisms of small RNA generation from cis-NATs in response to environmental and developmental cues. *Mol Plant*. 2013;6(3):704–15.
59. Creasey KM, Zhai J. miRNAs trigger widespread epigenetically-activated siRNAs from transposons in *Arabidopsis*. *Nature*. 2014;508:411–5.
60. Piriyapongsa J, Jordan IK. Dual coding of siRNAs and miRNAs by plant transposable elements. *Rna*. 2008;14(5):814–21.
61. Wang D, Qu Z, Yang L, Zhang Q, Liu ZH, Do T, et al. Transposable elements (TEs) contribute to stress-related long intergenic noncoding RNAs in plants. *Plant J Cell Mol Biol*. 2017;90(1):133–46.
62. Fowler S, Thomashow MF. *Arabidopsis* transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. *Plant Cell*. 2002;14(8):1675–90.
63. Zhang B, Pan X, Cannon CH, Cobb GP, Anderson TA. Conservation and divergence of plant microRNA genes. *Plant J*. 2006;46(2):243–59.
64. Pelaez P, Trejo MS, Iniguez LP, Estrada-Navarrete G, Covarrubias AA, Reyes JL, et al. Identification and characterization of microRNAs in *Phaseolus vulgaris* by high-throughput sequencing. *BMC Genomics*. 2012;13:83.
65. Bonnet E, Wuyts J, Rouze P, Van de Peer Y. Detection of 91 potential conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* identifies important target genes. *Proc Natl Acad Sci U S A*. 2004;101(31):11511–6.
66. Khaksefidi RE, Mirolohi S, Khalaji F, Fakhari Z, Shiran B, Ebrahimie E. Differential expression of seven conserved microRNAs in response to abiotic stress and their regulatory network in *Helianthus annuus*. *Front Plant Sci*. 2015;6:741.
67. Yu Y, Ni Z, Wang Y, Wan H, Hu Z, Jiang Q, et al. Overexpression of soybean miR169c confers increased drought stress sensitivity in transgenic *Arabidopsis thaliana*. *Plant Sci*. 2019;285:68–78.
68. Qin Z, Li C, Mao L, Wu L. Novel insights from non-conserved microRNAs in plants. *Front Plant Sci*. 2014;5:586.
69. Megha S, Basu U, Kav NNV. Regulation of low temperature stress in plants by microRNAs. *Plant Cell Environ*. 2018;41(1):1–15.
70. Lv DK, Bai X, Li Y, Ding XD, Ge Y, Cai H, et al. Profiling of cold-stress-responsive miRNAs in rice by microarrays. *Gene*. 2010;459(1–2):39–47.
71. Sunkar R, Zhu JK. Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. *Plant Cell*. 2004;16(8):2001–19.
72. Khraiweh B, Zhu JK, Zhu J. Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *Biochim Biophys Acta*. 2012;1819(2):137–48.
73. Dai X, Zhuang Z, Zhao PX. psRNA target: a plant small RNA target analysis server (2017 release). *Nucleic Acids Res*. 2018;46(W1):W49–54.
74. Kreps JA, Wu Y, Chang HS, Zhu T, Wang X, Harper JF. Transcriptome changes for *Arabidopsis* in response to salt, osmotic, and cold stress. *Plant Physiol*. 2002;130(4):2129–41.
75. Hahn A, Kilian J, Mohrholz A, Ladwig F, Peschke F, Dautel R, et al. Plant core environmental stress response genes are systemically coordinated during abiotic stresses. *Int J Mol Sci*. 2013;14(4):7617–41.
76. Bresso EG, Chorostecki U, Rodriguez RE, Palatnik JF, Schommer C. Spatial control of gene expression by miR319-regulated TCP transcription factors in leaf development. *Plant Physiol*. 2018;176(2):1694–708.
77. Soitamo AJ, Piippo M, Allahverdiyeva Y, Battchikova N, Aro EM. Light has a specific role in modulating *Arabidopsis* gene expression at low temperature. *BMC Plant Biol*. 2008;8:13.
78. Jiao X, Sherman BT, Huang da W, Stephens R, Baseler MW, Lane HC, et al. DAVID-WS: a stateful web service to facilitate gene/protein list analysis. *Bioinformatics*. 2012;28(13):1805–6.
79. Jin J, He K, Tang X, Li Z, Lv L, Zhao Y, et al. An *Arabidopsis* transcriptional regulatory map reveals distinct functional and evolutionary features of novel transcription factors. *Mol Biol Evol*. 2015;32(7):1767–73.
80. Palaniswamy SK, James S, Sun H, Lamb RS, Davuluri RV, Grotewold E. AGRIS and AtRegNet. A platform to link cis-regulatory elements and transcription factors into regulatory networks. *Plant Physiol*. 2006;140(3):818–29.
81. Tian F, Yang DC, Meng YQ, Jin J, Gao G. PlantRegMap: charting functional regulatory maps in plants. *Nucleic Acids Res*. 2019;48(D1):D1104–13.
82. Hartwell LH, Hopfield JJ, Leibler S, Murray AW. From molecular to modular cell biology. *Nature*. 1999;402(6761 Suppl):C47–52.
83. Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *J Stat Mech*. 2008;10:P10008.
84. Bouzroud S, Gouiaa S, Hu N, Bernadac A, Mila I, Bendou N, et al. Auxin response factors (ARFs) are potential mediators of auxin action in tomato response to biotic and abiotic stress (*Solanum lycopersicum*). *PLoS One*. 2018;13(2):e0193517.
85. BaozhuLi RF, Guo S, Wang P, Zhu X. The *Arabidopsis* MYB transcription factor, MYB111 modulates salt responses by regulating flavonoid biosynthesis. *Environmental and Experimental Botany*. 166. 103807. <https://doi.org/10.1016/j.envexpbot.2019.103807>.



86. Petridis A, Doll S, Nichelmann L, Bilger W, Mock HP. Arabidopsis thaliana G2-LIKE FLAVONOID REGULATOR and BRASSINOSTEROID ENHANCED EXPRESSION1 are low-temperature regulators of flavonoid accumulation. *New Phytol.* 2016;211(3):912–25.
87. Zhang X, Xia J, Liu YE. Genome-wide analysis of plant nat-siRNAs reveals insights into their distribution, biogenesis and function. *Genome Biol.* 2012;13:R20.
88. Wang H, Chung PJ, Liu J, Jang I-C, Kean MJ, Xu J, et al. Genome-wide identification of long noncoding natural antisense transcripts and their responses to light in *Arabidopsis*. *Genome Res.* 2014;24(3):444–53.
89. Jin H, Vasic V, Girke T, Lonardi S, Zhu JK. Small RNAs and the regulation of cis-natural antisense transcripts in *Arabidopsis*. *BMC Mol Biol.* 2008;9:6.
90. Tang G, Reinhart BJ, Bartel DP, Zamore PD. A biochemical framework for RNA silencing in plants. *Genes Dev.* 2003;17(1):49–63.
91. Martinez G, Choudury SG, Slotkin RK. tRNA-derived small RNAs target transposable element transcripts. *Nucleic Acids Res.* 2017;45(9):5142–52.
92. Cho J. Transposon-derived non-coding RNAs and their function in plants. *Front Plant Sci.* 2018;9:600.
93. Creasey KM, Zhai J, Borges F, Van Ex F, Regulski M, Meyers BC, et al. miRNAs trigger widespread epigenetically activated siRNAs from transposons in *Arabidopsis*. *Nature.* 2014;508(7496):411–5.
94. Loss-Morais G, Waterhouse PM, Margis R. Description of plant tRNA-derived RNA fragments (tRFs) associated with argonaute and identification of their putative targets. *Biol Direct.* 2013;8:6.
95. Ma L, Hong Z, Zhang Z. Perinuclear and nuclear envelope localizations of Arabidopsis ran proteins. *Plant Cell Rep.* 2007;26(8):1373–82.
96. Du J, Huang YP, Xi J, Cao MJ, Ni WS, Chen X, et al. Functional gene-mining for salt-tolerance genes with the power of *Arabidopsis*. *Plant J Cell Mol Biol.* 2008;56(4):653–64.
97. Leonhardt N, Kwak JM, Robert N, Waner D, Leonhardt G, Schroeder JI. Microarray expression analyses of *Arabidopsis* guard cells and isolation of a recessive abscisic acid hypersensitive protein phosphatase 2C mutant. *Plant Cell.* 2004;16(3):596–615.
98. Kaplan F, Guy CL. Beta-amylase induction and the protective role of maltose during temperature shock. *Plant Physiol.* 2004;135(3):1674–84.
99. Wang Z, Cui D, Liu J, Zhao J, Cheng L, Xin W, et al. Arabidopsis ZED1-related kinases mediate the temperature-sensitive intersection of immune response and growth homeostasis. *New Phytol.* 2017;215:711–24.
100. McCue AD, Slotkin RK. Transposable element small RNAs as regulators of gene expression. *Trends Genet.* 2012;28(12):616–23.
101. Howell MD, Fahlgren N, Chapman EJ, Cumbie JS, Sullivan CM, Givan SA, et al. Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/ DICER-LIKE4 pathway in *Arabidopsis* reveals dependency on miRNA- and tasiRNA-directed targeting. *Plant Cell.* 2007;19(3):926–42.
102. Binder S, Stoll K, Stoll B. P-class pentatricopeptide repeat proteins are required for efficient 5' end formation of plant mitochondrial transcripts. *RNA Biol.* 2013;10(9):1511–9.
103. Guo W-J, Ho T-HD. An abscisic acid-induced protein, HVA22, inhibits gibberellin-mediated programmed cell death in cereal aleurone cells. *Plant Physiol.* 2008;147(4):1710–22.
104. Samad AFA, Sajad M, Nazaruiddin N, Fauzi IA, Murad AMA, Zainal Z, et al. MicroRNA and transcription factor: key players in plant regulatory network. *Front Plant Sci.* 2017;8:565.
105. Atkinson NJ, Urwin PE. The interaction of plant biotic and abiotic stresses: from genes to the field. *J Exp Bot.* 2012;63(10):3523–43.
106. Confraria A, Martinho C, Elias A, Rubio-Somoza I, Baena-Gonzalez E. miRNAs mediate SnRK1-dependent energy signaling in *Arabidopsis*. *Front Plant Sci.* 2013;4:197.
107. Thiebaut F, Rojas CA, Almeida KL, Grativol C, Domiciano GC, Lamb CR, et al. Regulation of miR319 during cold stress in sugarcane. *Plant Cell Environ.* 2012;35(3):502–12.
108. Wang ST, Sun XL, Hoshino Y, Yu Y, Jia B, Sun ZW, et al. MicroRNA319 positively regulates cold tolerance by targeting OsPCF6 and OsTCP21 in rice (*Oryza sativa* L.). *PLoS One.* 2014;(3). <https://doi.org/10.1371/journal.pone.0091357>.
109. Ambawat S, Sharma P, Yadav NR, Yadav RC. MYB transcription factor genes as regulators for plant responses: an overview. *Physiol Mol Biol Plants.* 2013; 19(3):307–21.
110. Wang Y, Pang C, Li X, Hu Z, Lv Z, Zheng B, et al. Identification of tRNA nucleoside modification genes critical for stress response and development in rice and *Arabidopsis*. *BMC Plant Biol.* 2017;17(1):261.
111. Huber SM, Leonardi A, Dedon PC, Begley TJ. The versatile roles of the tRNA Epitranscriptome during cellular responses to toxic exposures and environmental stress. *Toxics.* 2019;7(1):17.
112. Taylor NL, Rudhe C, Hulett JM, Lithgow T, Glaser E, Day DA, et al. Environmental stresses inhibit and stimulate different protein import pathways in plant mitochondria. *FEBS Lett.* 2003;547(1–3):125–30.
113. Krimmer T, Rassow J, Kunau WH, Voos W, Pfanner N. Mitochondrial protein import motor: the ATPase domain of matrix Hsp70 is crucial for binding to Tim44, while the peptide binding domain and the Carboxy-terminal segment play a stimulatory role. *Mol Cell Biol.* 2000;20(16):5879–87.
114. Hong Ren M, Gray W. SAUR proteins as effectors of hormonal and environmental signals in plant growth. *Mol Plant.* 2015;8(8):1153–64.
115. Shi Y, Tian S, Hou L, Huang X, Zhang X, Guo H, et al. Ethylene signaling negatively regulates freezing tolerance by repressing expression of CBF and type-a ARR genes in *Arabidopsis*. *Plant Cell.* 2012;24(6):2578–95.
116. Masuda T. Recent overview of the mg branch of the tetrapyrrole biosynthesis leading to chlorophylls. *Photosynth Res.* 2008;96(2):121–43.
117. Zhang ZW, Yuan S, Xu F, Yang H, Chen YE, Yuan M, et al. Mg-protoporphyrin, haem and sugar signals double cellular total RNA against herbicide and high-light-derived oxidative stress. *Plant Cell Environ.* 2011; 34(6):1031–42.
118. Zhang Z-W, Wu Z-L, Feng L-Y, Dong L-H, Song A-J, Yuan M, et al. Mg-Protoporphyrin IX signals enhance Plant's tolerance to cold stress. *Front Plant Sci.* 2016;7:1545.
119. Sakuma Y, Maruyama K, Osakabe Y, Qin F, Seki M, Shinozaki K, et al. Functional analysis of an Arabidopsis transcription factor, DREB2A, involved in drought-responsive gene expression. *Plant Cell.* 2006;18(5):1292–309.
120. Hajzadeh M, Turkas M, Khawar KM, Unver T. miR408 overexpression causes increased drought tolerance in chickpea. *Gene.* 2015;555(2):186–93.
121. Jiang Q, Sun X, Niu F, Hu Z, Chen R, Zhang H. GmDREB1 overexpression affects the expression of microRNAs in GM wheat seeds. *PLoS One.* 2017;(5). <https://doi.org/10.1371/journal.pone.0175924>.
122. Ma C, Burd S. Lers a: miR408 is involved in abiotic stress responses in *Arabidopsis*. *Plant J.* 2015;84(1):169–87.
123. Kwon YJ, Park MJ, Kim SG, Baldwin IT, Park CM. Alternative splicing and nonsense-mediated decay of circadian clock genes under environmental stress conditions in *Arabidopsis*. *BMC Plant Biol.* 2014;14:136.
124. Zhang X, Gou M, Guo C, Yang H, Liu CJ. Down-regulation of Kelch domain-containing F-box protein in *Arabidopsis* enhances the production of (poly) phenols and tolerance to ultraviolet radiation. *Plant Physiol.* 2015;167(2):337–50.
125. Ma Z, Hu X, Cai W, Huang W, Zhou X, Luo Q, et al. Arabidopsis miR171-targeted scarecrow-like proteins bind to GT cis-elements and mediate gibberellin-regulated chlorophyll biosynthesis under light conditions. *PLoS Genet.* 2014. <https://doi.org/10.1371/journal.pgen.1004519>.
126. Wu G, Poethig RS. Temporal regulation of shoot development in *Arabidopsis thaliana* by miR156 and its target SPL3. *Development.* 2006; 133(18):3539–47.
127. Aukerman MJ, Sakai H. Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. *Plant Cell.* 2003; 15(11):2730–41.
128. Ding J, Shen J, Mao H, Xie W, Li X, Zhang Q. RNA-directed DNA methylation is involved in regulating photoperiod-sensitive male sterility in rice. *Mol Plant.* 2012;5(6):1210–6.
129. Hong JK, Choi HW, Hwang IS, Kim DS, Kim NH, Choi DS, et al. Function of a novel GDSL-type pepper lipase gene, CaGLIP1, in disease susceptibility and abiotic stress tolerance. *Planta.* 2008;227(3):539–58.
130. Gonzalez LE, Keller K, Chan KX, Gessel MM, Thines BC. Transcriptome analysis uncovers Arabidopsis F-BOX STRESS INDUCED 1 as a regulator of jasmonic acid and abscisic acid stress gene expression. *BMC Genomics.* 2017;18:533.
131. Li Q, Wang W, Wang W, Zhang G, Yang L, Wang Y, et al. Wheat F-box protein gene TaFBA1 is involved in plant tolerance to heat stress. *Front Plant Sci.* 2018;9:521.
132. Zhu L, Ow DW, Dong Z. Transfer RNA-derived small RNAs in plants. *Sci China Life Sci.* 2018;61(2):155–61.
133. Wang Q, Li T, Xu K, Zhang W, Wang X, Quan J, et al. The tRNA-derived small RNAs regulate gene expression through triggering sequence-specific degradation of target transcripts in the oomycete pathogen *Phytophthora sojae*. *Front Plant Sci.* 2016;7:1938.



134. Qin C, Xu PP, Zhang X, Zhang C, Liu CB, Yang DG, et al. Pathological significance of tRNA-derived small RNAs in neurological disorders. *Neural Regen Res*. 2020;15(2):212–21.
135. Cole C, Sobala A, Lu C, Thatcher SR, Bowman A, Brown JW, et al. Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *Rna*. 2009;15(12):2147–60.
136. Haussecker D, Huang Y, Lau A, Parameswaran P, Fire AZ, Kay MA. Human tRNA-derived small RNAs in the global regulation of RNA silencing. *Rna*. 2010;16(4):673–95.
137. Alves CS, Vicentini R, Duarte GT, Pinoti VF, Vincentz M, Nogueira FT. Genome-wide identification and characterization of tRNA-derived RNA fragments in land plants. *Plant Mol Biol*. 2017;93(1–2):35–48.
138. Yeung ML, Bennasser Y, Watashi K, Le SY, Houzet L, Jeang KT. Pyrosequencing of small non-coding RNAs in HIV-1 infected cells: evidence for the processing of a viral-cellular double-stranded RNA hybrid. *Nucleic Acids Res*. 2009;37(19):6575–86.
139. Garcia-Silva MR, Cabrera-Cabrera F, Guida MC, Cayota A. Hints of tRNA-derived small RNAs role in RNA silencing mechanisms. *Genes (Basel)*. 2012;3(4):603–14.
140. Kescu C, Kumar P, Kiran M, Su Z, Malik A, Dutta A. tRNA fragments (tRFs) guide ago to regulate gene expression post-transcriptionally in a Dicer-independent manner. *Rna*. 2018;24(8):1093–105.
141. Hackenberg M, Huang PJ, Huang CY, Shi BJ, Gustafson P, Langridge P. A comprehensive expression profile of microRNAs and other classes of non-coding small RNAs in barley under phosphorous-deficient and -sufficient conditions. *DNA Res*. 2013;20(2):109–25.
142. Hsieh LC, Lin SI, Shih AC, Chen JW, Lin WY, Tseng CY, et al. Uncovering small RNA-mediated responses to phosphate deficiency in *Arabidopsis* by deep sequencing. *Plant Physiol*. 2009;151(4):2120–32.
143. Ye R, Chen Z, Bi L, Jordan Rowley M, Xia N, Chai J, et al. A Dicer-independent route for biogenesis of siRNAs that direct DNA methylation in *Arabidopsis*. *Mol Cell*. 2016;61(2):222–35.
144. Xie M, Yu B. siRNA-directed DNA methylation in plants. *Curr Genomics*. 2015;16(1):23–31.
145. Wang Z, Wang F, Hong Y, Huang J, Shi H, Zhu JK. Two chloroplast proteins suppress drought resistance by affecting ROS production in guard cells. *Plant Physiol*. 2016;172(4):2491–503.
146. Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Cech M, et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res*. 2016;44(W1):W3–W10.
147. Cheng CY, Krishnakumar V, Chan AP, Thibaud-Nissen F, Schobel S, Town CD. Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *Plant J Cell Mol Biol*. 2017;89(4):789–804.
148. Axtell MJ. ShortStack: comprehensive annotation and quantification of small RNA genes. *Rna*. 2013;19(6):740–51.
149. Kramer MF. Stem-loop RT-qPCR for miRNAs. In: *Current protocols in molecular biology*; 2011. Chapter 15:Unit 15 10.
150. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>−</sup>(Delta Delta C(T)) method. *Methods*. 2001;25(4):402–8.
151. Jin J, Tian F, Yang DC, Meng YQ, Kong L, Luo J, et al. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res*. 2017;45(D1):D1040–5.
152. Rashmi KV, Gilad-Bachrach R. DART: dropouts meet multiple additive regression trees. 2015. Available online at <http://www.arxiv.org/pdf/150501866v1>.
153. Yilmaz A, Mejia-Guerra MK, Kurz K, Liang X, Welch L, Grotewold E. AGRIS: the *Arabidopsis* gene regulatory information server, an update. *Nucleic Acids Res*. 2011;39(Database issue):D1118–22.

# Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)



## Unpublished work: Functional analysis of miRNA overexpression lines in *Arabidopsis thaliana*

Kristin Habermann<sup>1</sup>, Oguz Top<sup>1</sup>, Wolfgang Frank<sup>1\*</sup>

1 Plant Molecular Cell Biology, Department Biology I, Ludwig-Maximilians-Universität München, LMU Biocenter, 82152 Planegg-Martinsried, Germany

### \*Correspondence

W. Frank, Plant Molecular Cell Biology, Department Biology I, Ludwig-Maximilians-Universität München, LMU Biocenter, Großhaderner Straße 2-4, 82152 Planegg-Martinsried, Germany. E-mail: wolfgang.frank@lmu.de. Telephone: +49 89 2180-74670

### **Abstract**

Micro RNAs (miRNAs) have an important impact under different stress conditions by regulating gene expression. Therefore, overexpression lines of different miRNAs have been analysed to uncover their role in plants in response to various environmental stimuli. Already available miRNA overexpression (OEX) lines of miR156h, miR398b and miR847 have been examined using a germination assay in response to salt, mannitol, abscisic acid (ABA), cold, heat and varying light conditions and have been compared to the *Arabidopsis thaliana* wild type (ecotype Columbia-0). Beside this, the miRNA levels have been quantified using RNA gel blot to confirm the overexpression of the respective *MIR* genes. Different miRNA targets have been predicted using the commonly available bioinformatic tool “psRNATarget” and quantitative reverse transcriptase PCR was used to quantify the mRNA levels of those miRNA targets. We could uncover seedlings of the miR847 OEX line to develop deviant phenotypes grown under salt (NaCl) and compared to the wild type. Further investigations on different NaCl concentrations have been performed to analyse the optimum concentration of NaCl where miR847 OEX line shows deviant phenotypes. Those phenotypes include tiny and dark green seedlings; seedlings with one or two pail cotyledonn

leaves, or seedlings with brown hypocotyls. miR847 OEX line and its putative mRNA target transcripts might play role in the development and/or stress acclimation.

### **Keywords**

Small non-coding RNA; ncRNA; gene regulation; *Arabidopsis thaliana*

### **Introduction**

Over the last years, there is rising evidence that small non-coding RNAs (sRNAs) with 20 to 24 nucleotides in length are important regulators of gene expression during the adaptive and protective responses of plants to various environmental stimuli (Manavella et al., 2019). sRNAs can execute diverse and complex gene regulation by different mechanisms. These mechanisms include epigenetic modifications to control the nuclear transcription (Holoch and Moazed, 2015, Bannister and Kouzarides, 2011, Khraiweh et al., 2010), mediated target mRNA cleavage and translational inhibition (Meister and Tuschl, 2004, Bartel, 2004, Kim, 2005). Based on their specific origin two major types of sRNAs can be distinguished as hairpin-derived sRNA (hpRNA) and small interfering RNA (siRNA) (Axtell, 2013). siRNAs are processed from double-stranded RNA (dsRNA) precursors, whereas hpRNAs have single-stranded precursors that fold back into a characteristic hairpin structure. microRNAs (miRNAs), a crucial class of hpRNA, regulate gene expression via sequence specific base pairing within mRNA molecules (Meyers et al., 2008). miRNAs are produced from characteristic stem-loop precursor transcripts encoded by *MIR* genes that are transcribed by RNA polymerase II (Pol II) generating a 5' capped and polyA-tailed primary miRNA transcript (pri-miRNA). These pri-miRNAs are further be processed into pre-miRNAs by DICER-LIKE1 (DCL1) enzymes with the help of accessory proteins HYPONASTIC LEAVES 1 (HYL1) and SERRATE (SE) (Park et al., 2002). A miRNA/miRNA\* duplex with 20-22 nucleotide in length is produced from the pre-miRNA by DCL1 and methylated through HUA ENHANCER 1 (HEN1), which is exported out of the nucleus into the cytoplasm by HASTY (HST) (Kim, 2004, Park et al., 2005). The mature miRNA of the miRNA/miRNA\* duplex is further loaded into the ARGONAUTE1 (AGO1) protein and combined to the RNA-

induced silencing complex (RISC). This complex guides the miRNA to its target RNA by sequence complementarity and mediates either target cleavage or translational inhibition (Wierzbicki et al., 2008, Voinnet, 2009). miRNAs control diverse biological processes including plant development, signal transduction, and response to environmental stress by regulating expression of many important genes (Palatnik et al., 2003, Sunkar and Zhu, 2004).

In recent years, several miRNAs as well as their targets could be identified to be important in the adaptation to environmental stress (Pegler et al., 2019, Shriram et al., 2016, Khraiwesh et al., 2012). For instance, miR408, one of the most conserved miRNAs in land plants, and their related RNA targets were previously identified to play a role in different stress conditions (Abdel-Ghany and Pilon, 2008, Zhang and Li, 2013, Thatcher et al., 2015, Axtell and Bowman, 2008). miR408 was analysed in response to different stresses and higher miRNA expression levels have been identified under cold, oxidative stress and salinity (Kantar et al., 2010, Li et al., 2011, Lu et al., 2005, Trindade et al., 2010, Zhang and Li, 2013, Thatcher et al., 2015, Abdel-Ghany and Pilon, 2008). In addition, miR408 targets are genes responsible for copper-binding proteins, which allows plants to better coordinate copper protein expression as well as development (Song et al., 2017).

Furthermore, in a previous study we have shown that sRNA classes especially miRNAs may act as regulators in retrograde signalling pathways (Habermann et al., 2020). *Arabidopsis thaliana gun1* and *gun5* mutants that show specific defects in retrograde signalling were used to analyse transcriptional changes of different sRNA classes as well as changes in mRNA and lncRNA levels in response to norflurazon (NF) (Habermann et al., 2020). We could detect a high number of differentially expressed miRNAs in response NF pointing to a considerable role of sRNAs in the adjustment of plastidic and nuclear gene expression caused by retrograde signals. The differentially expressed miRNAs that were identified in both *gun* mutants in response to NF included miRNAs showing the classical de-repression of nuclear genes encoding plastid proteins, which leads to the presumption that these miRNAs are controlled by retrograde signals and may have an impact on retrograde-controlled nuclear gene expression. Finally,

miRNA:RNA target pairs of interest could be detected that could act the adjustment of plastidic and nuclear gene expression or the adaptation of plastid function within retrograde signalling pathways (Habermann et al., 2020). For example, miR156h, miR398b and miR847 were found to be differentially regulated either in the NF treated wild type or in one of the *gun* mutants. The miR156 family is evolutionary conserved and involved in early root development, the control of flowering time as well as the control of different stages of the vegetative development by regulating mRNAs encoding important transcription factors including *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL)* (Wu et al., 2009, Wu and Poethig, 2006, Wang et al., 2009). Compared to wild type miR156h overexpression lines show an improved growth and survival rate under heat stress conditions, which is caused by decreased expression levels of its targets encoding SPL transcription factors (Stief et al., 2014). In contrast, miR398 was found to be downregulated by oxidative stress causing elevated levels of their cognate target RNAs *CSD1* and *CSD2* that code for a cytosolic and plastidic copper-zinc superoxide dismutase, respectively, that scavenge toxic superoxide radicals (Sunkar et al., 2006). miRNAs were also shown to play a role in hormone signalling. For example, elevated auxin levels lead to increased miR847 expression levels that cause concomitant downregulation of its cognate target transcript coding for *INDOLE-3-ACETIC ACID INDUCIBLE 28 (IAA28)*, a negative regulator of auxin signalling (Wang and Guo, 2015) and finally provokes lateral root development. Other miRNAs are important for the biotic stress response such as miR863-3p that has two major roles after infection of *Arabidopsis thaliana* with the bacterial pathogen *Pseudomonas syringae* (Niu et al., 2016). It was shown that miR863-3p mediates cleavage of transcripts encoding atypical receptor-like pseudokinases *ARLPK1* and *ARLPK2*, which leads to a defence response during the beginning of the infection, whereas at later stages miR863-3p forms a negative feedback loop by acting as a transcriptional inhibitor and reducing the expression levels of *SERRATE (SE)*, which is necessary to regulate the expression levels of miR863-3p.

In our previous study, we reported different sRNAs that are affected by retrograde signals (Habermann et al., 2020). Based on this, we selected different miRNAs

of interest and analysed different miRNA overexpression lines identify new potential roles of sRNAs under various stress conditions. One of these, miR847, seems to be involved in the response to abiotic stress, since the seedlings of the respective overexpression lines display specific phenotypic deviations in response to salt treatment.

## **Results**

### **Selection of miRNA overexpression lines**

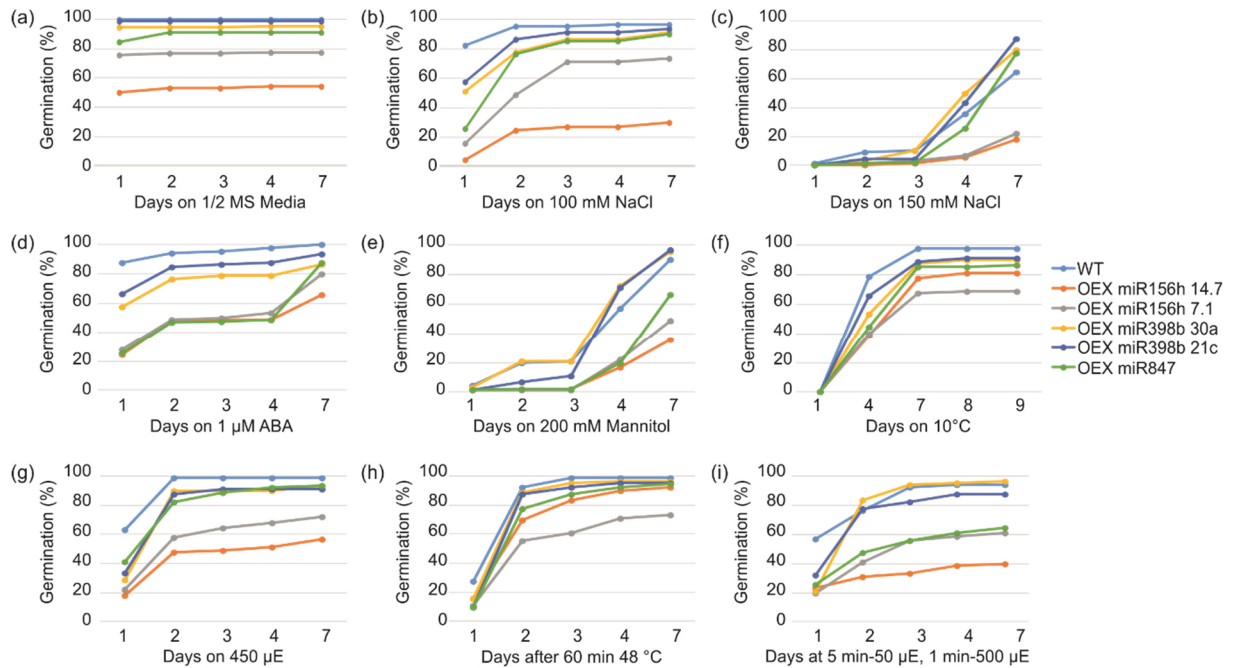
Recently, we reported that retrograde signals cause changes in the steady-state levels of different sRNAs classes (Habermann et al., 2020). Within the class of miRNAs an increase in response to NF in the *gun1* and *gun5* mutants was found to resemble the de-repression of nuclear encoded genes and it was assumed that these miRNAs may have an impact on retrograde-controlled nuclear gene expression. Besides, many differentially expressed mRNA transcripts were detected in response to NF suggesting mRNAs could regulate retrograde signalling and, more likely, they are themselves regulated by retrograde signals. miRNAs bind to their cognate targets by the recognition of reverse complementary sequences that was exploited to predict potential miRNA:mRNA target pairs that may act in the adjustment of plastidic and nuclear gene expression in retrograde signalling pathways (Habermann et al., 2020).

On the basis of miRNA and mRNA expression as well as biological function of miRNA targets, we selected miRNAs for functional analysis. Three miRNAs show differential expression (2-fold regulation, false discovery rate (FDR)  $\leq 0.05$ ) in response to retrograde signalling (Habermann et al., 2020). miR156h was identified to be 6.8-fold and 4.5-fold upregulated in both NF treated *gun1* and *gun5* mutants compared to the NF treated wild type, respectively. miR398b-5p was -3.2-fold downregulated in the NF treated wild type compared to the untreated wild type and further -2.5-fold downregulated in the NF treated *gun1* mutant compared to the NF treated wild type. miR847 was -5.2-fold downregulated in the NF treated wild type compared to the untreated wild type. Available miRNA overexpression (OEX) lines of miR156h (Stief et al., 2014), miR398b (Sunkar et al., 2006) and miR847 (Wang and Guo, 2015) were selected

for further analysis and compared to the *Arabidopsis thaliana* wild type (ecotype Columbia-0). Stief *et al.* (2014) reported miR156h overexpression lines to have an enhanced growth and survival rate under high temperatures caused by reduced expression levels of *SPL* transcription factors. Sunkar *et al.* (2006) reported miR398 to be decreased by oxidative stress leading to increased mRNA expression levels of *CSD1* and *CSD2* encoding for a cytosolic and plastidic copper-zinc superoxide dismutase, respectively. Wang and Guo (2015) show improved miR847 expression levels under enhanced auxin levels leading to decreased mRNA expression levels of *IAA28*, which regulates auxin signalling negatively causing lateral root development. In two of the three miRNA OEX lines, we obtained two independent lines from the same miRNA OEX lines: miR156h 14.7 and miR156h 7.1 as well as miR398b 30a and miR398b 21c. In addition, one OEX line of miR847 was used for further analysis. Those selected miRNA OEX lines have been analysed under various stress conditions to identify new potential roles of sRNAs.

### **Germination assay**

We used germination assay to improve insights about the specific role of different miRNAs. Therefore, different conditions were selected for the germination assay: salt (100 mM and 150 mM NaCl), mannitol (200 mM), abscisic acid (ABA, 1  $\mu$ M), cold (10°C), heat (48°C, 60 min) and varying light conditions (high light at 450  $\mu$ E and fluctuating light at 5 min - 50  $\mu$ E and 1 min - 500  $\mu$ E). The germination assay was performed on seeds from miRNA OEX lines and the wild type (Figure 1). Germination rates were scored on days 1, 2, 3, 4, and 7. In the case of cold treatment, germination was scored on days 1, 4, 7, 8, and 9.

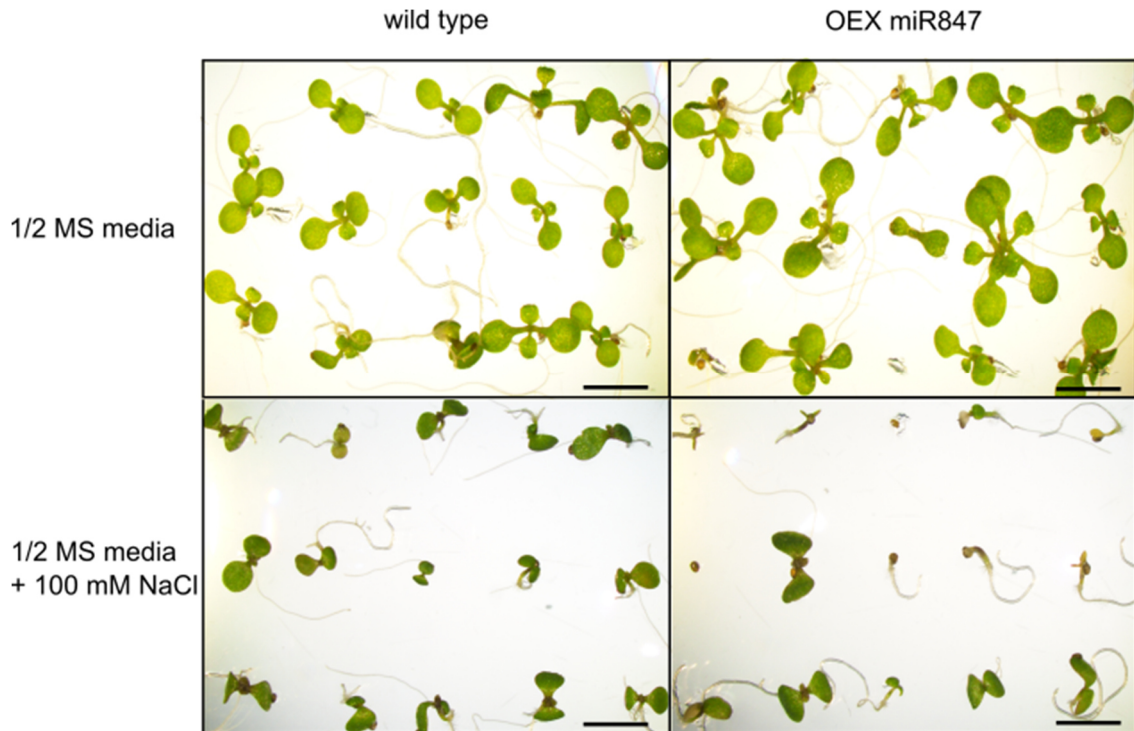


**Figure 1: Germination assay for different miRNA OEX lines and wild type) under different growth conditions. Germination rates were calculated for different treatments after two days of stratification and germination rates on ½ MS media served as control (a). Seeds were treated differently: 100 mM NaCl (b), 150 mM NaCl (c), 1 μM ABA (d), 200 mM mannitol (e), cold at 10°C (f), high light at 450 μE (g), heat for 60 min at 48°C (h) and fluctuating light at 5 min - 50 μE and 1 min - 500 μE (i). The emergence of the radicle was considered as successful germination and was determined microscopically.**

Germination frequencies of seeds from the different *MIR* OEX lines and WT control in various conditions were compared to their respective frequencies on ½ MS media. Differences in the germination of seeds were observed under control conditions where miR156h OEX remained below 80% of the germination rate after seven days. It is known that miR156 OEX lines show a delay in development and flowering, which might explain these germination differences (Schwab et al., 2005). Different NaCl concentrations like 100 mM (Figure 1b) and 150 mM NaCl (Figure 1c) reduced germination frequencies in some OEX lines. miR156h OEX showed only an approximately 50% germination rate after seven days under 100 mM NaCl as compared to the wild type. In general, higher salt concentration delayed seed germination and adversely affected seed germination frequencies, however, miR398b and miR847 OEX lines reached a higher germination rate than the wild type under 150 mM NaCl after seven days. The plant hormone ABA is known to inhibit seed germination and we observed delayed germination for all genotypes and the wild type in response to 1 μM ABA (Figure 1d). miR156h and



miR847 OEX lines show the lowest germination rates compared to the wild type. Mannitol is a sugar alcohol that is taken up by plant cells and causes oxidative and osmotic stress, that delays seed germination of all lines (Figure 1e). Seeds of the miR156h OEX lines (miR156h 14.7 and miR156h 7.1) showed less than 50% germination rates on 200 mM mannitol after seven days, whereas the miR398b OEX lines (miR398b 30a and miR398b 21c) showed nearly identical germination rates like the wild type. Under 10°C cold treatment a delay in germination was observed for all lines including the wild type and after 9 days nearly all lines reached the same germination rate as under control conditions (Figure 1f). 450  $\mu$ E were used for high light treatment and after seven days all lines including the wild type showed more than 50% of germination, but still miR156h OEX line displayed reduced germination compared to all other lines and the wild type (Figure 1g). After 60 min of heat stress at 48 °C, we observed a delay in the germination rate in all lines that was restored to control levels after seven days (Figure 1h). Under fluctuating light (5 min at 50  $\mu$ E and 1 min at 500  $\mu$ E) the different OEX lines and the wild type showed a large deviation between all the samples (Figure 1i). After seven days, only the wild type and miR398b OEX line reached germination rates higher than 80% of germination after seven days. The remaining OEX lines, miR156h and miR847, showed lower germination frequencies. miR847 OEX line had a germination rate of about 65%, miR156h 14.7 OEX genotype showed 40%, whereas the other genotype showed 60% seed germination rate after seven days under fluctuating light.

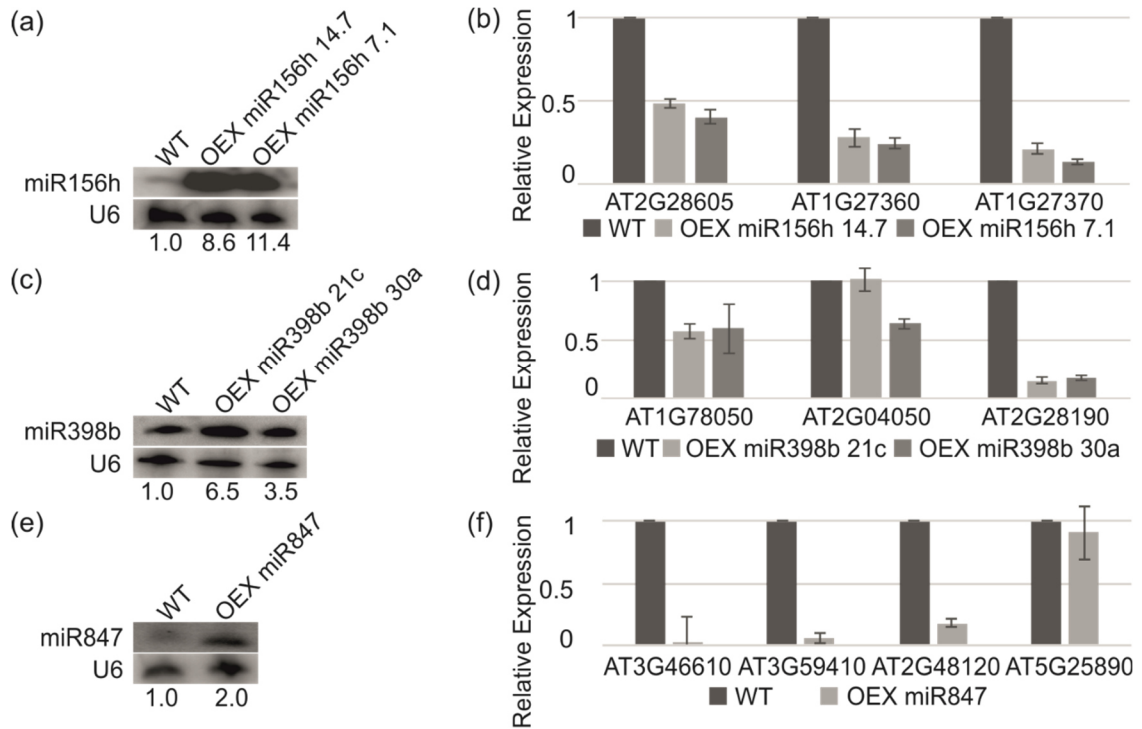


**Figure 2: Wild type and OEX miR847 line were grown on ½ MS medium supplemented with 100 mM NaCl plates seven days after stratification. The scale bars indicate 0.5 cm.**

The germination assay shows differences in all the OEX lines and genotypes. The wild type showed the highest germination frequencies under all conditions. Surprisingly, miR847 OEX lines showed pleiotropic phenotypes on ½ MS media + 100 mM NaCl. miR847 OEX seedlings developed tiny and dark green seedlings, seedlings with one or two pail cotyledons, or seedlings with brown hypocotyls under salt treatment.

### **miRNA:mRNA target analysis**

miRNAs are able to target mRNA transcripts and to mediate their cleavage. Previously, we reported miRNA target analysis, which was now used for additional investigations (Habermann et al., 2020). In addition, we performed RNA gel blots to confirm the overexpression of miRNAs in the OEX lines and qRT-PCR to address the impact on their target RNAs by comparing their steady-state levels between wild type and the miRNA OEX lines (Figure 3).



**Figure 3: Analysis of miRNA OEX lines and their cognate targets. RNA gel blots confirm the overexpression of miRNAs in different miRNA OEX lines: miR156h (a) miR398b (c) and miR847 (e). Signals were quantified and normalised to the constitutively expressed U6 snRNA and are indicated below the blots. Right panels (b, d and f) show graphs depicting the mRNA expression levels of predicted and validated target transcripts of the related miRNAs. Expression values were normalised to the constitutively expressed gene UBI1 (AT4G36800). Error bars indicate standard errors (n=3).**

miRNA156 is known to target mRNAs encoding SPL transcription factors, which are crucial for plant development (Stief et al., 2014). The RNA gel blots revealed a strong overexpression of miR156h in both independent miR156h OEX lines (Figure 3a). Further, we analysed the expression levels of two already known targets of miR156h encoding SPL transcription factors SPL10 (AT1G27360) and SPL11 (AT1G27370) and verified an efficient downregulation of these target transcripts. In addition, we analysed the transcript level of a predicted miR156h target encoding the plastid-localised PsbP domain-OEC23 like protein (AT2G28605) and detected a remarkable downregulation of this transcript in both independent miR156h OEX lines (Figure 3b). We also confirmed an overexpression of miR398b in both OEX lines (miR398b 30a and miR398b 21c) by RNA gel blot analysis (Figure 3c). Two interesting putative targets of miR398b encoding a plastid-localised phosphoglycerate/bisphosphoglycerate mutase (PGM, AT1G78050) and a multidrug and toxic compound extrusion efflux protein

(MATE, AT2G04050), respectively, were previously detected to be inversely correlated to the miRNA after norflurazon treatment (Habermann et al., 2020). We confirmed reduced steady-state levels for *PGM* in both miR398b OEX lines, whereas the transcript coding for the MATE transporter only showed decreased steady-state levels in the miR398b 30a OEX line (Figure 3d). The mRNA transcript coding for copper/zinc superoxide dismutase 2 (CSD2, AT2G28190) is an already validated miR398b target and operates as a positive control for the miR398b OEX line (Sunkar et al., 2006). The overexpression of miR847 was validated by RNA gel blot (Figure 3e). We could confirm a strong downregulation in the expression levels of three target transcripts of miR847 (Figure 3f). The first one encodes a plastid localised pentatricopeptide repeat (PPR) protein (AT3G46610), the second one encodes for the protein kinase general control non-repressible 2 (GCN2, AT3G59410) and the third one for the PALE CRESS (AT2G48120). However, we did not determine considerably decreased transcript level for the indole-3-acetic acid inducible 28 (IAA28, AT5G25890), which has been reported to be a target for miR847 (Wang and Guo, 2015).

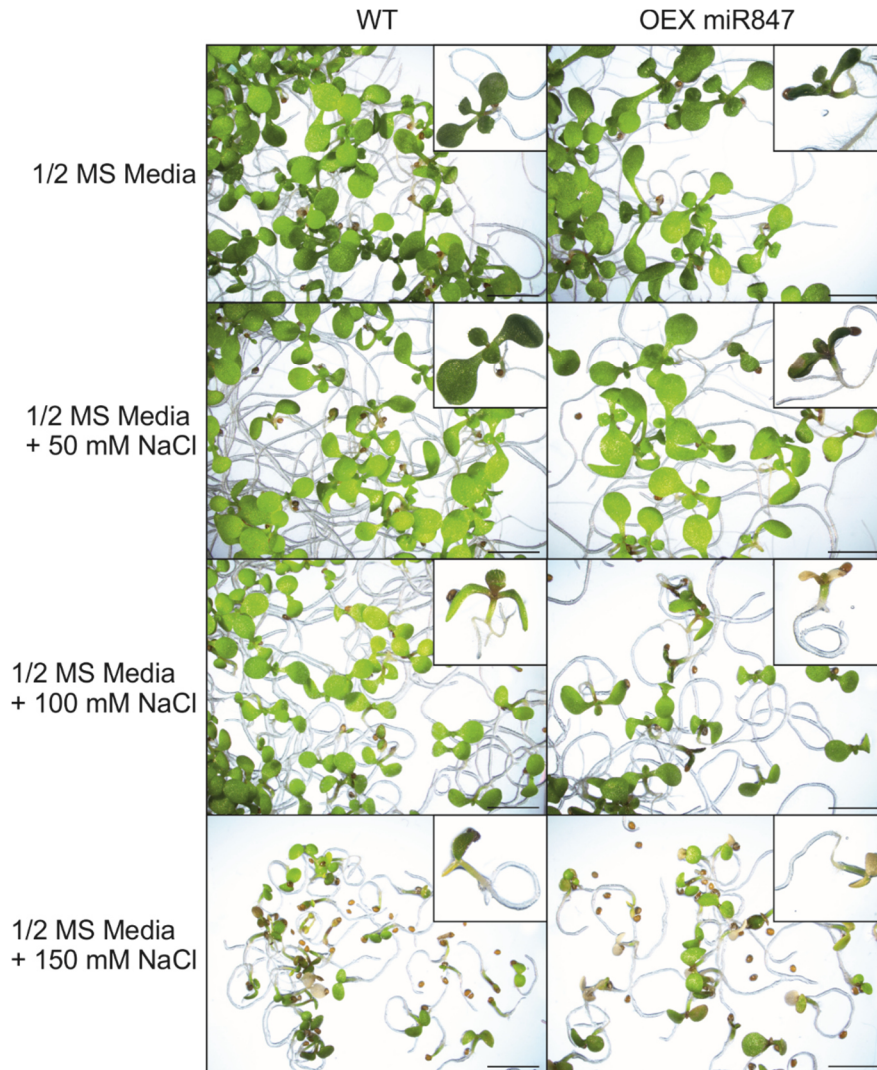
### Phenotypic analyses of miR847 OEX line

Seedlings of the miR847 OEX line were identified to show deviant phenotypes under salt, which require further analyses. The effect of salt concentration was investigated by seed germination assays for wild type and miR847 OEX lines, while using 114 seeds grown under different conditions: control ( $\frac{1}{2}$  MS media) and  $\frac{1}{2}$  MS media supplemented with 50 mM NaCl, 100 mM NaCl and 150 mM NaCl. After 3 and 7 days monitored the germination rates (Table 1). Under control conditions and 50 mM NaCl we observed less germination in miR847 OEX line compared to the wild type, but at least 90% of the seeds germinate. Higher salt concentrations delayed seed germinations and decreased seed germination in miR847 OEX line and the wild type. We observed a 10% decrease in the germination rate compared to the control conditions under 100 mM NaCl in miR847 OEX line and wild type. After three days of treatment a high difference between control conditions and 150 mM salt was observed with 11.5%. During later stages, after 7 days of germination, we detected lesser seed germination rate difference (2%) between control conditions and 150 mM salt.

**Table 1: Germination assay of wild type and miR847 OEX line under different conditions. In total, 114 seeds of wild type and miR847 OEX line were grown under different NaCl concentrations.  $\frac{1}{2}$  MS media was used and supplemented with different NaCl concentrations (50 mM NaCl, 100 mM NaCl, 150 mM NaCl) were used.**

Media	After 3 days		After 7 days	
	Wild type	OEX miR847	Wild type	OEX miR847
$\frac{1}{2}$ MS	96%	90%	99%	93%
$\frac{1}{2}$ MS + 50 mM NaCl	94.50%	88%	96%	90%
$\frac{1}{2}$ MS + 100 mM NaCl	87%	80%	87%	82%
$\frac{1}{2}$ MS + 150 mM NaCl	64.50%	53%	76%	74%

Plants, both wild type and miR847 OEX, germinated on  $\frac{1}{2}$  MS + 150 mM NaCl and showed delayed germination together with drastic phenotypic alterations including pale seedlings (Figure 4).

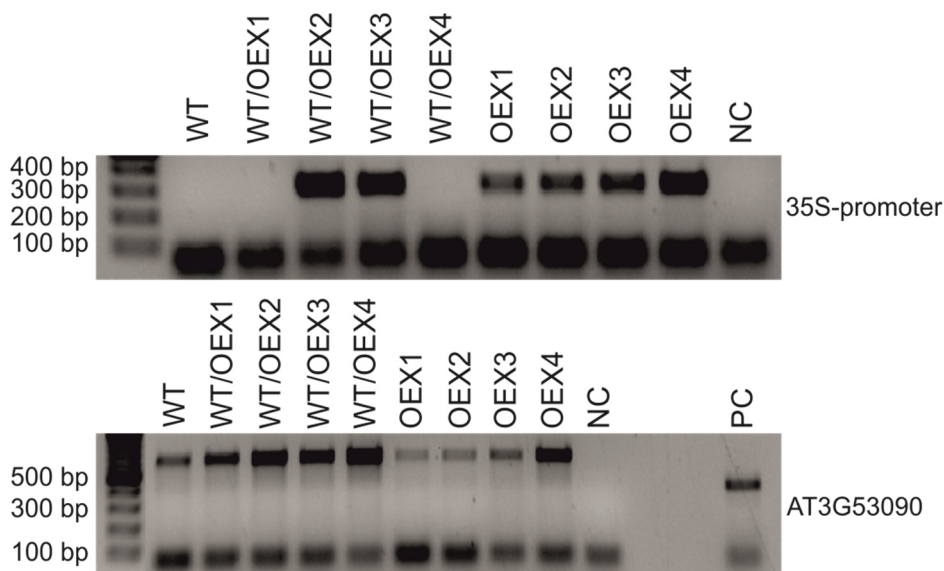


**Figure 4: Wild type (WT) and OEX miR847 seedlings were grown differently and monitored after seven days. The seedlings were grown on  $\frac{1}{2}$  MS media,  $\frac{1}{2}$  MS media + 50 mM NaCl,  $\frac{1}{2}$  MS media + 100 mM NaCl, and  $\frac{1}{2}$  MS media + 150 mM NaCl seven days after germination. The seedlings were examined microscopically with Nikon SMZ1500. The scale bar indicates 0.5 cm.**

Further analyses on miR847 OEX lines germinated on  $\frac{1}{2}$  MS media + 100 mM NaCl plates revealed that miR847 OEX seedlings had tiny and dark green cotyledons, seedlings with one or two pail cotyledons, or seedlings with brown hypocotyls (Figure 2 and 4). Despite these phenotypic differences, half of the miR847 OEX seedlings showed no altered phenotype compared to wild type. Therefore, miR847 OEX lines were further characterised on the molecular level by leaflet PCR in order to validate the presence of 35S promoter (Figure 5). A wild type seedling serves as a control. Further, we checked four seedlings showing wild type looking phenotypes (WT/OEX1, 2, 3, 4) as well as deviant



phenotypes: tiny and dark green seedling (OEX1), seedlings with two (OEX2) or one (OEX3) pail cotyledon leaves or seedlings showing a brown hypocotyl (OEX4) (Figure 5). The leaflet PCR indicates no 35S-promoter region for the wild type and two out of four wild type looking OEX miR847 seedlings (WT/OEX1 and WT/OEX4). Further, *UBI7* was used as a DNA control and it indicates that DNA could be extracted for all the samples.



**Figure 5:** Investigation on genomic level by leaflet PCR of 35S promoter of miR847 overexpression. Leaflet PCR indicate no 35S promoter in the wild type seedling (WT) and in two WT looking OEX miR847 seedlings (WT/OEX1 and WT/OEX4). 35S promoter was identified in two WT looking OEX miR847 seedlings (WT/OEX2 and WT/OEX3) and 4 different phenotypes (OEX1, OEX2, OEX3, OEX4). UBI7 could confirm DNA extraction for all samples. Water depicts as the negative control (NC) and positive control (PC) is validated cDNA from the wild type to prove the RT-PCR conditions.

We could identify OEX miR847 seedlings, which look like the wild type, but they do not have 35S promoter insertion of the overexpression of the *MIR847* gene, whereas all deviant phenotypes possess the 35S promoter insertion. This indicates a heterozygous OEX miR847 line.

## **Discussion**

Over the last years, many miRNAs have been identified to play an important role in plants under different stress conditions (Song et al., 2019). Furthermore, many investigations on miRNAs and their putative targets have been performed *in vitro* and *in silico* (Basso et al., 2019). For instance, miR398 was identified to be increased under heat stress to regulate its target *CSD2* and to better cope with

high temperatures (Fang et al., 2019). The *Arabidopsis thaliana gun* mutants show to have an influence on miRNA expression level under NF treatment, which was recently shown in our own study about the impact on sRNAs controlled by plastid-derived retrograde signals (Habermann et al., 2020). Both *gun1* and *gun5* mutants show changes in many miRNA expression steady-state levels in response to NF and compared to the NF treated *Arabidopsis thaliana* wild type, which leads an assumption that those miRNAs are regulators of the retrograde signalling by controlling transcript expression levels (Habermann et al., 2020). miRNA:mRNA pairs of interest were analysed regarding the correlation of their expression levels to assume a biological impact.

We selected three different miRNAs, miR156h, miR398b and miR847, which have been previously identified to be differentially expressed in retrograde signalling (Habermann et al., 2020). Available miRNA OEX lines have been used for further analyses. miR156 is known to control different vegetative stages of the plant as well as the flowering time mainly through regulation of the mRNA transcripts of the *SPL* transcription factors (Wu et al., 2009, Wu and Poethig, 2006, Wang et al., 2009). miR156 OEX lines could show a better survival under high temperatures by decreased mRNA transcript levels encoding *SPL* transcription factors (Stief et al., 2014). In contrast to miR156, miR398 was identified to be reduced under oxidative stress to maintain the mRNA transcripts of the *CSD2* to cope with very toxic superoxide (Sunkar et al., 2006). Auxin signalling can be controlled by miR847 and increased miRNA expression levels under auxin treatment cause a mediated cleavage of the *IAA28* transcript helping to regulate the lateral root development positively (Wang and Guo, 2015).

A germination assay under different conditions was performed to find more evidence on the different overexpression lines. Generally, miR156h OEX line showed a delay in germination in response to all stress conditions which is in line with previous investigations (Gao et al., 2017, Yu et al., 2015). Plants get adversely affected by salinity in terms of development and/or productivity, because they accumulate salt up to toxic levels leading to an excessive amount of reactive oxygen species (Miller et al., 2010, Fernando et al., 2018). This explains the delay in germination rates in all the OEX lines and the wild type under



salt stress. ABA and mannitol can affect the seed germination causing a delay as well as reduction in germination rates, which we could confirm in our germination assay (Yang et al., 2018). In addition, the cold treatment delayed germination in all lines, but all lines could reach the same germination rate after one week as in the control conditions. High and/or fluctuating light or heat stress did not cause any differences within the OEX lines. Heat and high light stresses affect the seeds only with regard to a delay in germination, but after seven days all lines, including the wild type and all OEX samples, are able to reach the same germination frequencies like under control conditions. Fluctuating light stress causes small differences between all the samples and after seven days miR398b OEX line could reach a slightly higher germination rate than the wild type. However, different germination was observed in different samples/conditions. Nevertheless, the wild type reached in nearly all conditions the highest germination rates. The miR847 OEX line evolve various phenotypes including tiny and dark green cotyledons, seedlings with one or two pail cotyledons, or seedlings with brown hypocotyl on  $\frac{1}{2}$  MS media containing 100 mM NaCl.

We could validate the miRNA overexpression of the related *MIR* genes in all OEX lines under normal growth conditions via RNA gel blot. Furthermore, the expression levels of the miRNA targets have been quantified for each OEX line. Both well-known miR156 targets, *SPL10* and *SPL11*, show decreased mRNA expression levels. Further, the putative target transcript encoding for a plastid-localised PsbP domain-OEC23 like protein was reduced by 50% compared to the wild type, which leads to the assumption that this target might be regulated through miR156h. The PsbP domain belongs to one of the extrinsic proteins of photosystem II and it is required for the stabilisation and activation of the cofactors of the oxygen-evolving complex (OEC) (Gregor and Britt, 2000). Consequently, miR156h is able to regulate transcripts, which are important for different development stages as well as the flowering time. miR156h mediate the cleavage of the transcript of a plastid-localised PsbP domain-OEC23 like protein, which is essential for the control of the Photosystem II. The validated miR398b target, *CSD2*, could confirm decreased steady-state levels in miR398b OEX line compared to the wild type. In addition, two new putative targets have been

checked regarding their transcript expression levels, but the mRNA transcript encoding for the MATE protein suggests no cleavage mediated through miR398b, whereas the second putative target shows reduced expression values compared to the control. The beginning of the starch biosynthesis is mediated by the enzyme PGM, which converts glucose-6-phosphate into glucose-1-phosphate (Zhang et al., 2019). The *Arabidopsis thaliana pgm* mutant is defective in starch biogenesis leading to adverse effects on the cell wall metabolism and structure (Engelsdorf et al., 2017). Taken together, reduced transcript levels of *PGM* in miR398b OEX line indicate a mediated cleavage of the *PGM* transcript by miR398b. Unfortunately, we could not detect downregulation of the *IAA28* mRNA transcript targeted by miR847, but all three newly predicted targets showed decreased steady-state levels of their mRNA transcripts compared to the wild type. The PPR protein known as *LOW PHOTOSYNTHETIC EFFICIENCY 1 (LPE1)* binds to the mRNA of *psbA* that encodes the D1 protein of the photosystem II reaction centre (Jin et al., 2018). Therefore, miR847 seems to be involved in the photosynthesis by targeting *LPE1*. The GCN2 kinase, a universal regulator, is required for the translation initiation factor eIF2 $\alpha$ , which is necessary in stress signal transduction or different immune responses (Lokdarshi et al., 2020). The PALE CRESS (PAC) protein is important for the leaf as well as for chloroplast development and its albinotic *pac* mutant possess drastic translational deficiencies demonstrating the importance of this protein in plastid gene expression (Meurer et al., 2017). Consequently, miR847 might target some very important mRNA target transcripts, that play role in the development and/or stress acclimation.

In addition, the miR847 OEX line was checked with regard to different NaCl concentrations. We could identify different phenotypes in the miR847 OEX line under 100 mM NaCl: tiny and dark green cotyledons, seedlings with one or two pail cotyledons, or seedlings with brown hypocotyls. However, half of the NaCl treated seedlings did not show a phenotypic change in the miR847 OEX line compared to the NaCl treated wild type. Therefore, different miR847 OEX seedlings were analysed concerning the presence of the 35S promoter region. We could not detect the 35S promoter region for the overexpression of the

*MIR847* gene for the wild type looking miR847 OEX seedlings indicating a heterozygous miR847 OEX line.

Here, we could demonstrate that miRNAs can play important roles in response to different stress conditions. miR847 OEX line evolve pleiotropic phenotypes in response to NaCl treatment compared to the wild type. Nevertheless, genotypic analyses of the *MIR847* promotor region have been performed indicating a heterozygous miR847 OEX line. Therefore, further investigations are required to uncover the role of miR847 generally and under NaCl treatment. Firstly, it is indispensable to create a homozygous miR847 OEX line through generation of a new miR847 OEX line. Secondly, the NaCl treatment has to be repeated to validate the current findings and to do additional analyses like quantification of the miRNA target transcripts, which might include also the quantification of the protein levels compared to the wild type and in response to salt.

## **Experimental procedures**

### **Plant material and germination assay conditions**

*Arabidopsis thaliana* wild type (ecotype Columbia-0) and the overexpression (OEX) mutants, OEX miR156h (Stief et al., 2014), OEX miR398b (Sunkar et al., 2006), OEX miR847 (Wang and Guo, 2015) and OEX miR863 (Niu et al., 2016), were used for this study. 30 surface-sterilised seeds from the different OEX lines and wild type were grown on one plate with three biological replicates to monitor germination. The overall germination rate was determined on ½ Murashige and Skoog (MS) medium containing 1.5% sucrose. Further, seeds were treated differently with: salt (100 mM NaCl and 150 mM NaCl), Absciscic acid (1 µM ABA), mannitol (200 mM mannitol), cold (10°C), high light (450 µE), heat (60 min at 48°C) and fluctuating light (5 min-50 µE, 1 min-500 µE). Germination was monitored microscopically.

In addition, 38 surface-sterilised seeds from OEX miR847 and wild type were grown on one plate and in three biological replicates to monitor germination at ½ MS media containing 1.5% sucrose and ½ MS media supplemented with 50 mM NaCl, 100 mM NaCl and 150 mM NaCl.

### **Growth conditions**

Surface-sterilised seeds were incubated on ½ MS agar plates containing 1.5% sucrose. In addition, ½ MS agar plates containing 100 mM NaCl were used. After vernalisation (2 days at 4°C in darkness) the seeds were grown for one week at 22°C and 110 µE. Whole plants were harvested and immediately frozen in liquid nitrogen and stored at -80°C until RNA isolation. All experiments were performed in three biological replicates for each genotype.

### **RNA isolation**

RNA isolation was performed by using TRIzol reagent (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's protocol. Agarose gel electrophoresis was used to monitor the RNA integrity. The RNA concentration and purity were evaluated by spectrophotometry (260 nm/280 nm and 260 nm/230 nm absorbance ratios) (NanoDrop™ 2000/2000c Spectrophotometer, Thermo Fisher Scientific, Waltham, MA, USA).

### **RNA gel blot**

RNA gel blot was modified from Pall and Hamilton (2008). In total, 30 µg RNA was loaded on a 15% PAGE and run for 3 h at 60 V, electroblotted on Amersham Hybond-N neutral Membrane (GE Healthcare UK Limited, Buckinghamshire, UK) at 180 mA for 1 h and the RNA was crosslinked with 1-Ethyl-3-[3-dimethylaminopropyl] carbodiimide hydrochloride (EDC) for 1 h at 60°C.

The detection of specific RNAs was performed according to probe preparation and hybridization conditions from Pall and Hamilton (2008). The membrane was pre-hybridised in Church-Buffer (1 M Na<sub>2</sub>HPO<sub>4</sub>, 1 M NaH<sub>2</sub>PO<sub>4</sub>, 0.5 M EDTA, pH 8.0, 20 % SDS, 20 x SSC, 100x Denhardt's reagent) for at least 2 h at hybridisation temperature. DNA oligonucleotides (Table S1) complementary to miRNA were radiolabeled with [<sup>32</sup>P]γ-ATP (Hartmann Analytic GmbH, Braunschweig, Germany) using T4 Kinase (New England Biolabs, Ipswich, MA, USA) according to the manufacturer's protocol. After heat inactivation (68°C, 10 min) the probes and membranes were hybridised overnight and washed according to Pall and Hamilton (2008). The membranes were incubated on a phosphor imaging screen, scanned after overnight incubation using Typhoon

TRIO (GE Healthcare Typhoon) and U6 normalised by using Quantity One Software (Bio-Rad, Hercules, CA, USA).

### Quantitative RT-PCR

To remove residual genomic DNA RNA samples were treated with DNase I (New England Biolabs, Ipswich, MA, USA) to perform cDNA synthesis. In total, 2 µg RNA was incubated with DNaseI (2 U, New England Biolabs, Ipswich, MA, USA) for 30 min at 37°C. 2.5 µl 50 mM EDTA were added and incubated for 10 min at 65°C to inactivate the DNaseI. The RNA, 100 pmol of an oligo-dT23VN oligonucleotide and 10 mM dNTPs were denatured at 65°C for 5 min and transferred to ice. Afterwards, M-MuLV reverse transcriptase (200 U, New England Biolabs, Ipswich, MA, USA) was used according to the manufactures protocol. We performed reverse transcription (RT) PCR to examine effective cDNA synthesis by the use of the gene *UBI7* (AT3G53090) (Table S2).

20 ng/µl cDNA were used for each qRT-PCR together with gene-specific primers and an EvaGreen qPCR mix. First, the samples were pre-heated for 2 min at 95°C. 40 qRT-PCR cycles were performed with these conditions: 12 sec at 95°C, 30 sec at 58°C and 15 sec at 72°C in three technical triplicates with the CFX Connect Real-Time PCR device (Bio-Rad, Hercules, CA, USA).

$2^{-\Delta\Delta ct}$  method was used to calculate changes in gene expression with the help of the  $\Delta ct$ -values (Livak and Schmittgen, 2001). Normalisation was done by using the housekeeping gene *UBI1* (AT4G36800). All oligonucleotide sequences which were used in this study are listed in Table S2.

### DNA extraction

Genomic DNA extraction was performed according to Edwards et al. (1991). Approximately 100 mg of plant material was disrupted together with 200 µl plant DNA extraction buffer (200 mM Tris - HCl pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5% SDS) and a metal beat (3 mm Ø) with TissueLyser II (QIAGEN GmbH, Hilden, Germany) for five times for 30 seconds at 30 movements /s and incubated for 5 min at room temperature. After 7 min of centrifugation at 13,000 g, 130 µl of the supernatant and 150 µl isopropanol (-20°C) were incubated for 2 min at room temperature. DNA was pelleted by 10 min centrifugation at 4°C at 13,000 g. The

pellet was washed with 500 µl 70% ethanol (-20°C) and centrifuged for 5 min at 13,000 g. The dried DNA pellet was resuspended in 30 µl TE Buffer (10 mM Tris pH 8, 1 mM EDTA) with RNase A (5 mg/mL).

Leaflet RT-PCR was performed using gene-specific primers and the gene *UBI7* (AT3G53090) as a positive control (Table S2). The RT-PCR was examined by agarose gel electrophoresis.

### **Accession numbers**

ATG accession numbers: MIR156h (AT5G55835), MIR398b (AT5G14545), MIR847 (AT1G07051), GUN1 (AT2G31400), GUN5 (AT5G13630), plastid-localised PsbP domain-OEC23 like protein (AT2G28605), MATE efflux family protein (AT2G04050), SPL10 (AT1G27360), SPL11 (AT1G27370), PGM (AT1G78050), CSD2 (AT2G28190), PPR protein (AT3G46610), GCN2 (AT3G59410), PALE CRESS (AT2G48120), IAA28 (AT5G25890), UBI7 (AT3G53090), UBI1 (AT4G36800)

### **Acknowledgments**

This work was supported by the German Research Foundation (SFB-TRR 175, grants to W.F. project C03).

### **Conflict of Interest**

The authors declare no conflict of interest.

### **Author contributions**

WF and KH designed the research; KH performed the research; KH, OT and WF analysed the data and wrote the paper. All authors read and approved the final manuscript.

### **Supporting Information**

Supporting tables:

**Table S1: Sequences of DNA oligonucleotides used for RNA gel blot in this study.**

Probe Name	Sequence
ath-miR156h	GTGCTCTCTTTCTTCTGTCA
ath-miR398b-5p	GTGTGTTCTCATATCAACCCT
ath-miR847	CATCAAGAAGAAGAGGAGTGA
ath-miR863-3p	ATTATGTCTTGTGCTCTCAA

**Table S2: Sequences of oligonucleotides used for quantitative RT-PCR this study.**

Primer Name	Sequence forward Primer	Sequence reverse Primer
AT3G53090	ACTCAAACCCGGTGGGAAGG	AGGGCACAATCTGCCGGTTA
AT4G36800	CTGTTACGGAACCCAATTC	GGAAAAAGGTCTGACCGACA
AT2G28605	CGACTCGCAATCAGAACGAAG	TTGCCTAGTTAACATTCTTTGTGGA
AT1G27360	TCATGGCGAAGATGTGGGAGAA	AAATGGCTGCACATGTGGTTGG
AT1G27370	TGGGTTCTCAGCAGGGAAATCC	GGAGTGTGTTTGATCCCTTGTGA
AT1G78050	TGGGAGCTAGAGAATGAAAAAG A	TAAACGCCTCCTTCAACAACC
AT2G04050	CCACAAGGCATAGGAGCCAG	ACACATACTCTGCAACTGCC
AT2G28190	ACACACGGAGCTCCAGAAGA	TGTTGTTTCTGCCACGCCAT
AT3G46610	AGATTGAAGCCTGCGGTAGC	TCATAGCACCCAAAAGGCTGT
AT3G59410	GATTGGCAATGGAGATTGTTGC	CTTTAACCTCTCCAACCTTCACTCTC
AT5G25890	CCGGAGAAACCTAACGGCAC	TGGAACCTCTTCCATGTTGATCT
AT2G48120	ACGATGATGAATGGCAAGCTC	CCTCGTGCCAATTCACGGT
35S-promoter	GGAAAGGCTATCGTTCAAGATGCCT	TACCCTGCTCAACTCATGTCACCC



## References

- ABDEL-GHANY, S. E. & PILON, M. 2008. MicroRNA-mediated systemic down-regulation of copper protein expression in response to low copper availability in Arabidopsis. *J Biol Chem*, 283, 15932-15945.
- AXTELL, M. J. 2013. Classification and comparison of small RNAs from plants. *Annu Rev Plant Biol*, 64, 137-159.
- AXTELL, M. J. & BOWMAN, J. L. 2008. Evolution of plant microRNAs and their targets. *Trends Plant Sci*, 13, 343-349.
- BANNISTER, A. J. & KOUZARIDES, T. 2011. Regulation of chromatin by histone modifications. *Cell Res*, 21, 381-395.
- BARTEL, D. P. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 116, 281-297.
- BASSO, M. F., FERREIRA, P. C. G., KOBAYASHI, A. K., HARMON, F. G., NEPOMUCENO, A. L., MOLINARI, H. B. C. & GROSSI-DE-SA, M. F. 2019. MicroRNAs and new biotechnological tools for its modulation and improving stress tolerance in plants. *Plant Biotechnol J*, 17, 1482-1500.
- EDWARDS, K., JOHNSTONE, C. & THOMPSON, C. 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res*, 19, 1349.
- ENGELSDORF, T., WILL, C., HOFMANN, J., SCHMITT, C., MERRITT, B. B., RIEGER, L., FRENGER, M. S., MARSCHALL, A., FRANKE, R. B., PATTATHIL, S. & VOLL, L. M. 2017. Cell wall composition and penetration resistance against the fungal pathogen *Colletotrichum higginsianum* are affected by impaired starch turnover in Arabidopsis mutants. *J Exp Bot*, 68, 701-713.
- FANG, X., ZHAO, G., ZHANG, S., LI, Y., GU, H., LI, Y., ZHAO, Q. & QI, Y. 2019. Chloroplast-to-nucleus signaling regulates microRNA biogenesis in Arabidopsis. *Dev Cell*, 48, 371-382 e4.
- FERNANDO, V. C. D., AL KHATEEB, W., BELMONTE, M. F. & SCHROEDER, D. F. 2018. Role of Arabidopsis ABF1/3/4 during det1 germination in salt and osmotic stress conditions. *Plant Mol Biol*, 97, 149-163.
- GAO, R., WANG, Y., GRUBER, M. Y. & HANNOUFA, A. 2017. miR156/SPL10 modulates lateral root development, branching and leaf morphology in Arabidopsis by silencing AGAMOUS-LIKE 79. *Front Plant Sci*, 8, 2226.
- GREGOR, W. & BRITT, R. D. 2000. Nitrogen ligation to the manganese cluster of photosystem II in the absence of the extrinsic proteins and as a function of pH. *Photosynth Res*, 65, 175-185.
- HABERMANN, K., TIWARI, B., KRANTZ, M., ADLER, S. O., KLIPP, E., ARIF, M. A. & FRANK, W. 2020. Identification of small non-coding RNAs responsive to GUN1 and GUN5 related retrograde signals in Arabidopsis thaliana. *Plant J*, 104, 138-155.
- HOLOCH, D. & MOAZED, D. 2015. RNA-mediated epigenetic regulation of gene expression. *Nat Rev Genet*, 16, 71-84.
- JIN, H., FU, M., DUAN, Z., DUAN, S., LI, M., DONG, X., LIU, B., FENG, D., WANG, J., PENG, L. & WANG, H. B. 2018. LOW PHOTOSYNTHETIC EFFICIENCY 1 is required for light-regulated photosystem II biogenesis in Arabidopsis. *Proc Natl Acad Sci U S A*, 115, E6075-E6084.

- KANTAR, M., UNVER, T. & BUDAK, H. 2010. Regulation of barley miRNAs upon dehydration stress correlated with target gene expression. *Funct Integr Genomics*, 10, 493-507.
- KHRAIWESH, B., ARIF, M. A., SEUMEL, G. I., OSSOWSKI, S., WEIGEL, D., RESKI, R. & FRANK, W. 2010. Transcriptional control of gene expression by microRNAs. *Cell*, 140, 111-122.
- KHRAIWESH, B., ZHU, J. K. & ZHU, J. 2012. Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *Biochim Biophys Acta*, 1819, 137-148.
- KIM, V. N. 2004. MicroRNA precursors in motion: exportin-5 mediates their nuclear export. *Trends Cell Biol*, 14, 156-159.
- KIM, V. N. 2005. Small RNAs: classification, biogenesis, and function. *Mol Cells*, 19, 1-15.
- LI, T., LI, H., ZHANG, Y. X. & LIU, J. Y. 2011. Identification and analysis of seven H<sub>2</sub>O<sub>2</sub>-responsive miRNAs and 32 new miRNAs in the seedlings of rice (*Oryza sativa* L. ssp. indica). *Nucleic Acids Res*, 39, 2821-2833.
- LIVAK, K. J. & SCHMITTGEN, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-ΔΔC<sub>T</sub></sup> Method. *Methods*, 25, 402-408.
- LOKDARSHI, A., GUAN, J., URQUIDI CAMACHO, R. A., CHO, S. K., MORGAN, P. W., LEONARD, M., SHIMONO, M., DAY, B. & VON ARNIM, A. G. 2020. Light activates the translational regulatory kinase GCN2 via reactive oxygen species emanating from the chloroplast. *Plant Cell*, 32, 1161-1178.
- LU, S., SUN, Y. H., SHI, R., CLARK, C., LI, L. & CHIANG, V. L. 2005. Novel and mechanical stress-responsive microRNAs in *Populus trichocarpa* that are absent from *Arabidopsis*. *Plant Cell*, 17, 2186-2203.
- MANAVELLA, P. A., YANG, S. W. & PALATNIK, J. 2019. Keep calm and carry on: miRNA biogenesis under stress. *Plant J*, 99, 832-843.
- MEISTER, G. & TUSCHL, T. 2004. Mechanisms of gene silencing by double-stranded RNA. *Nature*, 431, 343-349.
- MEURER, J., SCHMID, L. M., STOPPEL, R., LEISTER, D., BRACHMANN, A. & MANAVSKI, N. 2017. PALE CRESS binds to plastid RNAs and facilitates the biogenesis of the 50S ribosomal subunit. *Plant J*, 92, 400-413.
- MEYERS, B. C., AXTELL, M. J., BARTEL, B., BARTEL, D. P., BAULCOMBE, D., BOWMAN, J. L., CAO, X., CARRINGTON, J. C., CHEN, X., GREEN, P. J., GRIFFITHS-JONES, S., JACOBSEN, S. E., MALLORY, A. C., MARTIENSSSEN, R. A., POETHIG, R. S., QI, Y., VAUCHERET, H., VOINNET, O., WATANABE, Y., WEIGEL, D. & ZHU, J. K. 2008. Criteria for annotation of plant microRNAs. *Plant Cell*, 20, 3186-3190.
- MILLER, G., SUZUKI, N., CIFTCI-YILMAZ, S. & MITTLER, R. 2010. Reactive oxygen species homeostasis and signalling during drought and salinity stresses. *Plant Cell Environ*, 33, 453-467.
- NIU, D., LII, Y. E., CHELLAPPAN, P., LEI, L., PERALTA, K., JIANG, C., GUO, J., COAKER, G. & JIN, H. 2016. miRNA863-3p sequentially targets negative immune regulator ARLPKs and positive regulator SERRATE upon bacterial infection. *Nat Commun*, 7, 11324.

- PALATNIK, J. F., ALLEN, E., WU, X., SCHOMMER, C., SCHWAB, R., CARRINGTON, J. C. & WEIGEL, D. 2003. Control of leaf morphogenesis by microRNAs. *Nature*, 425, 257-263.
- PALL, G. S. & HAMILTON, A. J. 2008. Improved northern blot method for enhanced detection of small RNA. *Nat Protoc*, 3, 1077-1084.
- PARK, M. Y., WU, G., GONZALEZ-SULSER, A., VAUCHERET, H. & POETHIG, R. S. 2005. Nuclear processing and export of microRNAs in Arabidopsis. *Proc Natl Acad Sci U S A*, 102, 3691-3696.
- PARK, W., LI, J., SONG, R., MESSING, J. & CHEN, X. 2002. CARPEL FACTORY, a dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. *Curr Biol*, 12, 1484-1495.
- PEGLER, J. L., OULTRAM, J. M. J., GROF, C. P. L. & EAMENS, A. L. 2019. Profiling the abiotic stress responsive microRNA landscape of Arabidopsis thaliana. *Plants (Basel)*, 8, 58.
- SCHWAB, R., PALATNIK, J. F., RIESTER, M., SCHOMMER, C., SCHMID, M. & WEIGEL, D. 2005. Specific effects of microRNAs on the plant transcriptome. *Dev Cell*, 8, 517-527.
- SHRIRAM, V., KUMAR, V., DEVARUMATH, R. M., KHARE, T. S. & WANI, S. H. 2016. MicroRNAs as potential targets for abiotic stress tolerance in plants. *Front Plant Sci*, 7, 817.
- SONG, X., LI, Y., CAO, X. & QI, Y. 2019. MicroRNAs and their regulatory roles in plant-environment interactions. *Annu Rev Plant Biol*, 70, 489-525.
- SONG, Z., ZHANG, L., WANG, Y., LI, H., LI, S., ZHAO, H. & ZHANG, H. 2017. Constitutive expression of miR408 improves biomass and seed yield in Arabidopsis. *Front Plant Sci*, 8, 2114.
- STIEF, A., ALTMANN, S., HOFFMANN, K., PANT, B. D., SCHEIBLE, W. R. & BAURLE, I. 2014. Arabidopsis miR156 regulates tolerance to recurring environmental stress through SPL transcription factors. *Plant Cell*, 26, 1792-1807.
- SUNKAR, R., KAPOOR, A. & ZHU, J. K. 2006. Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in Arabidopsis is mediated by downregulation of miR398 and important for oxidative stress tolerance. *Plant Cell*, 18, 2051-2065.
- SUNKAR, R. & ZHU, J. K. 2004. Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. *Plant Cell*, 16, 2001-2019.
- THATCHER, S. R., BURD, S., WRIGHT, C., LERS, A. & GREEN, P. J. 2015. Differential expression of miRNAs and their target genes in senescing leaves and siliques: insights from deep sequencing of small RNAs and cleaved target RNAs. *Plant Cell Environ*, 38, 188-200.
- TRINDADE, I., CAPITAO, C., DALMAY, T., FEVEREIRO, M. P. & SANTOS, D. M. 2010. miR398 and miR408 are up-regulated in response to water deficit in Medicago truncatula. *Planta*, 231, 705-716.
- VOINNET, O. 2009. Origin, biogenesis, and activity of plant microRNAs. *Cell*, 136, 669-687.
- WANG, J. J. & GUO, H. S. 2015. Cleavage of INDOLE-3-ACETIC ACID INDUCIBLE28 mRNA by microRNA847 upregulates auxin signaling to modulate cell proliferation and lateral organ growth in Arabidopsis. *Plant Cell*, 27, 574-590.

- WANG, J. W., CZECH, B. & WEIGEL, D. 2009. miR156-regulated SPL transcription factors define an endogenous flowering pathway in *Arabidopsis thaliana*. *Cell*, 138, 738-749.
- WIERZBICKI, A. T., HAAG, J. R. & PIKAARD, C. S. 2008. Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell*, 135, 635-648.
- WU, G., PARK, M. Y., CONWAY, S. R., WANG, J. W., WEIGEL, D. & POETHIG, R. S. 2009. The sequential action of miR156 and miR172 regulates developmental timing in *Arabidopsis*. *Cell*, 138, 750-759.
- WU, G. & POETHIG, R. S. 2006. Temporal regulation of shoot development in *Arabidopsis thaliana* by miR156 and its target SPL3. *Development*, 133, 3539-3547.
- YANG, B., SONG, Z., LI, C., JIANG, J., ZHOU, Y., WANG, R., WANG, Q., NI, C., LIANG, Q., CHEN, H. & FAN, L. M. 2018. RSM1, an *Arabidopsis* MYB protein, interacts with HY5/HYH to modulate seed germination and seedling development in response to abscisic acid and salinity. *PLoS Genet*, 14, e1007839.
- YU, N., NIU, Q. W., NG, K. H. & CHUA, N. H. 2015. The role of miR156/SPLs modules in *Arabidopsis* lateral root development. *Plant J*, 83, 673-685.
- ZHANG, H. & LI, L. 2013. SQUAMOSA promoter binding protein-like7 regulated microRNA408 is required for vegetative development in *Arabidopsis*. *Plant J*, 74, 98-109.
- ZHANG, Y., HE, P., MA, X., YANG, Z., PANG, C., YU, J., WANG, G., FRIML, J. & XIAO, G. 2019. Auxin-mediated statolith production for root gravitropism. *New Phytol*, 224, 761-774.

## **DISCUSSION**

Until now, few next generation sequencing (NGS) studies have been performed on sRNAs, which also take into account the expression levels of putative mRNA target transcripts. The aim of my thesis was to investigate the role of sRNA as well as their targets in *Arabidopsis thaliana* in retrograde signalling and during cold conditions using high-throughput sequencing techniques. Furthermore, functional analysis experiments have been performed to confirm the sRNA, lncRNA and mRNA sequencing data. In the first publication, mRNA/lncRNA sequencing was performed and combined with sRNA sequencing to better understand the function of sRNAs and lncRNAs in retrograde signalling (Habermann et al., 2020). Seedlings of *Arabidopsis thaliana* wild type and the two retrograde signalling mutants, *gun1* and *gun5*, were treated with and without NF for four days, and the differential expression of different RNA classes were analysed with focus on retrograde-dependent expression level changes. In the second publication, sRNA sequencing was performed with 2-weeks old *Arabidopsis thaliana* wild type seedlings after 3 h, 6 h and 2 d of cold treatment (4°C) (Tiwari et al., 2020) and published mRNA transcriptome data was used to complete the analyses (Garcia-Molina et al., 2020) to identify the effect of cold-responsive regulation through sRNAs by RNA interference.

Little is known about the role of lncRNAs and sRNAs in plastid-to-nucleus signalling. Until now, few transcriptomic data approaches analysed the retrograde signalling *gun* mutants under NF treatment, however, they neglected to analyse lncRNAs and sRNAs. Previous *Arabidopsis thaliana* microarray studies have been performed on seedlings of wild type as well as of *gun1*, *gun1-9*, *gun2* and *gun5* mutants to gain more knowledge about the retrograde signalling pathway (Koussevitzky et al., 2007). An overlap of 43% and 67% could be observed for the up- and downregulated differentially expressed genes (DEGs) in the wild type in response to NF between our study and Koussevitzky *et al.* (2007) (Publication I, Figure S9a, b). Additionally, 56% and 50% of the identified DEGs in our study detected in the NF-treated *gun1-9* and *gun5* compared to the treated wild type,

respectively, were also detected by Koussevitzky *et al.* (2007). New methodologies have been developed leading to replacement of microarray experiments by NGS tools with a better sequencing quality and depth. Zhao *et al.* (2019b) performed RNA-sequencing with seedlings of *Arabidopsis thaliana* wild type and *gun1-9* mutants treated with/without NF to discover accessory roles of *GUN1*. In addition, recently, NGS was used to study the regulation of genes involved in anthocyanin biosynthesis in response to NF in *Arabidopsis thaliana* wild type seedlings as well as *gun1-1*, *gun4-1* and *gun5-1* mutant seedlings (Richter *et al.*, 2020). A quite good overlap could be observed between both NGS studies and our study. In detail, 55% and 49% of our DEGs could also be identified to be differentially expressed in Richter *et al.* (2020) in the NF-treated *gun1* and *gun5* mutant compared to the NF-treated wild type, respectively. An even higher overlap of DEGs of about 65% could be observed in our NF-treated *gun1* mutant compared to the NF-treated wild type to Zhao *et al.* (2019b). Consequently, we noticed a high accordance between our study and other transcriptome studies, although different methods, growth conditions and mutants were used.

Our transcriptome data could confirm the classical de-repression of *PhANGs* under NF in both retrograde signalling mutants. Generally, we identified a large overlap between the nucleus encoded DEGs in NF-treated *gun1* and *gun5* mutant compared to the NF-treated wild type; but interestingly, all the plastid encoded DEGs are downregulated in the NF-treated *gun1* mutant compared to the NF-treated wild type, whereas the plastid encoded DEGs are upregulated in the NF-treated *gun5* mutant compared to the NF-treated wild type. Previously, Woodson *et al.* (2013) suggest the sigma factors (*SIG2*, *SIG6*) and the plastid encoded RNA polymerase (PEP) are able to control the plastid gene transcription by retrograde signalling networks. It can be hypothesised that PEP is activated through *GUN1* and in the *gun1* mutant the PEP is perturbed, which might cause the changes in the plastid encoded DEGs compared to the wild type in response to NF.

Besides mRNA data, lncRNA data was also analysed to identify an impact of ncRNA on plastid-derived signals. One lncRNA (AT4G13495) showed the

classical de-repression like *PhANGs* (2-fold regulation (*gun* NF/WT NF), false discovery rate ( $FDR \leq 0.05$ ) in treated *gun1* and *gun5* mutant compared to the NF-treated wild type. Interestingly, three different miRNA precursors (*MIR5026*, *MIR850* and *MIR863*) overlap in sense direction with this lncRNA, and all sRNAs generated from this locus have been identified to be differentially expressed in at least one sample in response to NF (downregulated in WT NF/WT and upregulated in *gun* NF/WT NF; Publication I, Table 2). Further, the same AT4G13495 lncRNA transcript could be detected to be downregulated after 2 d of cold treatment (-2,3 fold,  $FDR \leq 0.05$ ), but no differentially expressed sRNAs could be observed for any time point (3 h, 6 h, or 2 d). Nevertheless, it can be suggested that miRNAs can either be processed from the lncRNA itself or from the individual miRNA precursor transcripts. To predict putative miRNA targets we used psRNATarget tool with optimised stringent parameters. We could not detect mRNA targets for miR5026, whereas two targets, coding for a chloroplast RNA binding protein (AT1G09340) and a threonine-tRNA ligase (AT2G04842), could be predicted for miR850. The miRNA and both targets have been detected to be upregulated in the NF-treated *gun5* mutant compared to the treated wild type. Because miRNAs mediate cleavage of mRNA transcripts and miR850 and both putative transcripts are upregulated in expression levels it leads to the assumption that miR850 does not cleave those transcripts. Additionally, miR863 could be detected to be upregulated (fold change  $\geq +2$ ;  $FDR \leq 0.05$ ) in both NF-treated *gun* mutants compared to the treated wild type, even if the changes in expression levels of the predicted *SERRATE* (*SE*) transcript was not detected to be significant ( $FDR \leq 0.05$ ) in those samples (Publication I, Table 2). SE protein is known to be a regulator of the miRNA biogenesis pathway, which affects the processing of various miRNAs (Meng et al., 2012).

We demonstrate in Publication I and II the role of non-coding sRNAs in retrograde signalling pathway and cold acclimation, since comprehensive changes in the steady-state levels could be observed for different sRNA classes (miRNAs, nat-siRNAs and other sRNA producing loci). In the first Publication, we detected generally more downregulated sRNAs in the NF-treated wild type compared to the untreated control and more upregulated sRNAs in both treated *gun* mutants



compared to the treated wild type (Habermann et al., 2020). Thus, increased sRNA processing in the *gun* mutants in response to NF resembles the de-repression of nuclear encoded genes, which leads to the hypothesis that sRNAs can play a role in regulation of retrograde-controlled nuclear gene expression. In the second Publication, an overall reduction of different sRNA classes was observed over the time course in response to low temperatures. However, this could also be due to general changes of the levels of main components required for sRNA biogenesis. To elucidate this case, the expression levels of the transcripts of those components, such as *HEN1*, *RDR6*, *DCL1*, *HST*, *HYL1*, *SE* and *SGS3*, have been analysed and we could see that they are unchanged over the time course of cold treatment. Consequently, we can assume that sRNA processing is reduced due to a reduction of the transcripts of the sRNA precursors.

The highest changes of steady-state levels of miRNAs were identified in the NF-treated wild type compared to the untreated control with the highest downregulation of miR169g-3p (-151.6-fold; FDR  $\leq$  0.05). The same miRNA is the most upregulated miRNA found in the NF-treated *gun5* mutant compared to the NF-treated wild type (38.2-fold; FDR  $\leq$  0.05). The miR169 family was already detected to be a drought, heat and salt responsive miRNA in different plants (Li et al., 2008, Yang et al., 2019, Shukla et al., 2018, Xu et al., 2014). Under cold acclimation, the highest changes of the expression levels could be observed after 6 h of cold treatment. miR395e was detected to be the most upregulated miRNA (53.0-fold; FDR  $\leq$  0.05) and miR447b the most downregulated miRNA (-28.0-fold; FDR  $\leq$  0.05). The gene family of miR395 can be encoded from six different *Arabidopsis thaliana* loci and those individual members have been identified to play different roles in abiotic stress conditions like salt or mannitol, because single nucleotide differences in the miRNA family members cause changes in the miRNA targets (Kim et al., 2010).

Since miRNAs are powerful regulators of RNA interference to control gene expression, we performed miRNA target prediction and associated our sRNA sequencing data with the mRNA data to identify miRNA:RNA pairs. In total, we discovered 47 differentially expressed miRNAs belonging to 119 non-redundant

putative targets in response to NF and further 93 differentially expressed miRNAs belonging to 352 non-redundant putative targets in response to cold treatment. In total, we detected 16 miRNAs to be differentially expressed in response to NF as well as under cold treatment.

The most conserved miRNAs in *Arabidopsis thaliana* comprised miR156 and miR157 (Reinhart et al., 2002, Wu and Poethig, 2006, He et al., 2018) and we detected both miRNAs in response to NF as well as in cold acclimation. Under NF treatment, miR156 isoforms have been detected to be downregulated in the treated wild type compared to the untreated wild type (fold change  $\leq -2$ ; FDR  $\leq 0.05$ ) and to be upregulated in both treated *gun* mutants compared to the treated wild type (fold change  $\geq +2$ ; FDR  $\leq 0.05$ ). miR157a was only detected to be downregulated in the NF-treated *gun5* mutant compared to the treated wild type (miR157a-3p: -48.8-fold; miR157a-5p: -6.67; FDR  $\leq 0.05$ ). In addition, both miRNAs, miR156 and miR157, were upregulated after 2 d of cold treatment (miR156f-3p: 5,61-fold; miR156f-5p: 4,56-fold; miR157a-3p: 4,55-fold; miR157b-3p: 33,7-fold; miR157c-3p: 3,2-fold; FDR  $\leq 0.05$ ). After miRNA target prediction, a high number of targets could be determined and it was found that different mRNAs encoding transcription factors of the *SPL* family are targeted by miRNAs. In response to NF, the *SPL* targets have the same correlation in expression levels like the miRNAs. So, miR156 and *SPL10* transcript target were upregulated in the NF-treated *gun5* mutant compared to the treated wild type (2.0-fold; FDR  $\leq 0.05$ ). Nevertheless, both miRNAs, miR156 and miR157, were upregulated after 2 d of cold treatment, whereas its putative target *SPL3* was detected to be downregulated at the same time point (-2.9-fold; FDR  $\leq 0.05$ ), showing the classical anticorrelation between the miRNA and its target. However, a similar regulation of the miRNA and its target can still denote a regulation, since the miRNA binds to protein complexes leading to increased target transcription or no cleavage caused by variations in the miRNA or target sequences (Baulina et al., 2016, Saito and Saetrom, 2012). The *SPL* transcription factors are known to regulate the early root development, flowering time as well as other different stages of the vegetative development, showing the importance of those putative targets as well as of the regulation through miRNAs under different stress

conditions (Ren and Tang, 2012, Cui et al., 2014, Stief et al., 2014, Gao et al., 2017).

Furthermore, we detected miR159 isoforms to be downregulated in the NF-treated *gun1* mutant compared to the NF-treated wild type (miR159b-5p: -3,75-fold;  $FDR \leq 0.05$ ), but upregulated after 3 h of low temperature (miR159c: 2,14-fold;  $FDR \leq 0.05$ ) and downregulated after 2 d of cold acclimation (miR159a: -2,55-fold; miR159b-3p: -2,17-fold;  $FDR \leq 0.05$ ). The RNA target prediction, the retrograde signalling study could not detect an informative putative target, since the target was not differentially expressed, but the cold study could determine a mRNA target encoding a mitochondrial translocase TIM44 related protein to be anticorrelated to miR159. This result suggests that under low temperatures an altered protein import due to miR159-mediated regulation of its target, which is in line with the assumption that environmental stress conditions can repress and stimulate protein import (Taylor et al., 2003, Lister et al., 2004). As an example, TIM44 together with heat shock protein 70 (HSP70) promote the protein import from the inner membrane to the mitochondrial matrix (Krimmer et al., 2000, Murcha et al., 2003).

In addition, miR395b and miR395c were detected to be downregulated in the NF-treated wild type compared to the untreated control (miR395b: -5.48-fold; miR395c: -5.35-fold;  $FDR \leq 0.05$ ) and upregulated in the treated *gun5* mutant compared to the NF-treated wild type (miR395b: 6.36-fold;  $FDR \leq 0.05$ ). Moreover, miR395c was discovered to be downregulated after 6 h of cold treatment (miR395c: -2.12-fold;  $FDR \leq 0.05$ ). Fang *et al.* (2018) reports that tocopheroles are involved in retrograde signalling and miR395 mediated cleavage of *APS* mRNA transcript. Fang *et al.* (2018) identified tocopheroles to positively regulate PAP accumulation, which has a positive effect on miRNA biogenesis: PAP negatively regulates XRN2, which in turn regulates mRNA and pri-miRNA levels by degrading 5' uncapped mRNA. In addition, the mRNA transcript encoding the enzyme APS was validated to be a target for miR395 (Liang et al., 2010) and APS is catalysing the initial step in PAP synthesis (Klein and Papenbrock, 2004). Our findings are in agreement with the previous assumption by Fang *et al.* (2018) that reduced miR395b levels in the NF-treated

wild type lead to increased *APS* transcript levels. Increased *APS* levels cause enhanced PAP synthesis that in turn acts as a retrograde inhibitor of XRN5 leading to enhanced pri-miRNA and mature miRNA levels. Furthermore, miR398 was detected to be increased under high temperatures to regulate its target *CSD2* resulting in a better stress adaptation under high temperatures (Fang et al., 2018). We could detect another interesting mRNA targeted by miR395, which codes for the magnesium-chelatase subunit *GUN5* (AT5G13630), an enzyme involved in the chlorophyll biosynthesis pathway catalysing the insertion of  $Mg^{2+}$  into protoporphyrin IX. The *GUN5* mRNA transcript levels were detected to be downregulated in the NF-treated wild type compared to the untreated control (-2.36-fold;  $FDR \leq 0.05$ ), upregulated in the treated *gun5* mutant compared to the NF-treated wild type (4.03-fold;  $FDR \leq 0.05$ ) and upregulated after 6 h of cold treatment (2.2-fold;  $FDR \leq 0.05$ ). The *gun5* mutant has a single nucleotide substitution in the *GUN5* gene, which causes a defective magnesium-chelatase. Although the expression levels of the miRNA:mRNA pairs are not anticorrelated in response to NF, the elevated miRNA expression levels can maybe balance an increased transcription rate of its target mRNA to maintain physiologically related steady-state levels. In the NF-treated wild type, the *GUN5* mRNA transcript is decreased due to NF-triggered retrograde signalling, because lacking of the retrograde signalling in the *gun5* mutant the mRNA transcript levels remain increased, which cannot be efficiently decreased by enhanced miRNA levels. A study from Zhang *et al.* (2011) showed that Mg-Proto-IX-derived signals cause an induction of Alternative oxidase 1a (AOX1a). The enzyme AOX1a catalyses the reduction from  $O_2$  to  $H_2O$ , which emit excess energy as heat. Furthermore, the Mg-Proto-IX-derived signals lead to enhanced activities of antioxidant enzymes, which serve as preservation of redox equilibrium under cold stress conditions (Zhang et al., 2016).

Besides miRNAs, other sRNA classes have been analysed with regard to the differential regulation of their expression levels. Fewer differentially expressed sRNAs derived from lncRNAs have been identified in both publications. Nevertheless, we could detect an even higher number of differentially regulated nat-siRNAs under cold acclimation as well as in response to NF. In numbers, 326

non-redundant *cis*-NATs and 142 non-redundant *trans*-NATs pairs have been detected to produce differentially expressed siRNAs under cold acclimation, whereas 73 non-redundant *cis*-NATs and 193 non-redundant *trans*-NATs pairs have been detected to produce differentially expressed siRNAs in response to NF. In both publications, nat-siRNAs with less than five normalised reads have been removed to ensure that we detect abundant sRNAs. It is noteworthy to mention that more *cis*-nat-siRNAs were detected to be differentially regulated under low temperatures and more differentially expressed *trans*-nat-siRNAs were detected in response to NF. An overlap between nat-siRNAs in both publications could be identified, which leads to the assumption that also nat-siRNAs play a role in stress response. In response to NF, a huge number nuclear encoded differentially expressed genes extend the knowledge about genes, which can be controlled by retrograde signals. Under low temperatures a predominance of pre-tRNA or transposable element (TE) serves as a non-coding partner of the gene pairs.

To deepen our knowledge about the function of miRNAs and their corresponding mRNA targets, we selected differentially expressed miRNAs for further analysis (Unpublished work: Functional analysis of miRNA overexpression lines in *Arabidopsis thaliana*). We examined putative mRNA targets to identify new potential roles under various stress conditions.

Since both *Arabidopsis thaliana* retrograde signalling defective mutants, *gun1* and *gun5*, as well as the wild type exhibit changes in the miRNA and mRNA expression steady-state levels under NF, we hypothesized that miRNAs can have an impact on the retrograde signalling by controlling the expression levels of their associated mRNA target. Three different miRNAs, miR156h, miR398b and miR847, and their putative targets were selected for functional analysis experiments. miR156 belongs to the evolutionary highly conserved miRNA family in plants and is involved in controlling early root development, flowering time and other vegetative stages by regulating mRNA transcripts encoding *SPL* transcription factors (Gao et al., 2017, Wang, 2014, Bari et al., 2013). Stief *et al.*

(2014) studied miR156h overexpression lines and described better growth as well as enhanced survival rates in response to high temperatures compared to the *Arabidopsis thaliana* wild type. miR156h overexpression causes decreased transcript levels of the *SPL* transcription factor genes by increased miRNA cleavage that results in a better adaptation to heat stress conditions (Stief et al., 2014). Conversely, in response to oxidative stress, decreased miR398 levels lead to an increase of their validated transcript targets encoding cytosolic CSD1 and plastidic CSD2 protein (Sunkar et al., 2006). Both dismutases are required to protect plant cells against the toxic superoxide. Further, *INDOLE-3-ACETIC ACID INDUCIBLE 28 (IAA28)* is a negative regulator of auxin signalling by repressing ARFs. Therefore, mRNA transcript levels of *IAA28* needs to be reduced under auxin treatment, which is mediated by miR847 and results in an increased lateral root development (Wang and Guo, 2015).

We performed germination assays using miR156h, miR398b and miR847 overexpression lines under control (1/2 MS media, 22°C, long day) and different stress conditions such as high salinity (100 mM NaCl, 150 mM NaCl), cold temperature (10°C), mannitol (200 mM), ABA (1 µM), high temperature (48°C for 60 min), high light (450 µE) or fluctuating light (5 min - 50 µE and 1 min - 500 µE) to gain more knowledge about the specific role of miRNAs. The overexpression line of miR156h is already known to have a delay in germination (Gao et al., 2017, Yu et al., 2015), a feature that we could also verify in all tested conditions. In response to salt stress, we observed a delay in the germination rate for all overexpression lines as well as for the wild type, which can be explained by an accumulation of salt that leads to increased ROS levels (Fernando et al., 2018). Additionally, ABA and mannitol is known to impair the seed germination rates in the wild type (Yang et al., 2018a). Under low temperatures, the plant lines show initially a delay in germination compared to the control conditions, but after one week all lines reached the same germination rate as under the control conditions without differences in the lines. This germination delay in cold is an already known phenomenon, since low temperatures affect plant metabolism, growth and development (Sasaki et al., 2015). High light or heat stress could not cause an effect on the different overexpression lines. Only the overexpression line

miR398b had a slightly increased germination rate compared to the wild type under fluctuating light stress conditions. Nevertheless, the wild type exhibited except of one (fluctuating light) the highest germination rates in nearly all stress conditions and we did not detect phenotypical differences between the overexpression lines and the wild type with one exception. Seedlings of the overexpression line miR847 showed various phenotypes on the  $\frac{1}{2}$  MS media containing 100 mM NaCl. Those phenotypes include tiny and dark green seedlings, seedlings with brown hypocotyls or seedlings with one or two pale cotyledon leaves. Based on these findings, further investigations have been performed. The growth behaviour of the *MIR847* overexpression line has been analysed at different NaCl concentrations (50 mM, 100 mM and 150 mM). We could identify different phenotypes in the *MIR847* overexpression line upon salt stress ( $\frac{1}{2}$  MS media containing 100 mM NaCl). Nevertheless, half of the miR847 overexpression line seedlings treated with NaCl showed a wild type phenotype. The verification of the presence of the 35S promoter region of the *MIR847* overexpression construct revealed that this construct is not present in the wild type looking plants. Taken together, the miR847 overexpression line is a heterozygous line, because we identified wild type looking plants without 35S promoter region of the *MIR847* and phenotypes containing the 35S promoter region of the *MIR847*. Further, a homozygous line has to be generated for the miR847 overexpression line to perform further analysis. Using homozygous miR847 overexpression line, miR847 mRNA targets can be studied and validated to identify the importance of this miRNA during salt treatment or other stress conditions.

Besides the germination assay, the *MIR* genes related to the specific overexpression lines have also been confirmed. The RNA gel blot was used to detect an increase in expression levels of the associated miRNA in the overexpression lines compared to the wild type. In addition, we checked the mRNA expression levels of putative and already validated miRNA targets by quantitative RT-PCR. Besides the two validated targets of miR156h, *SPL10* and *SPL11*, also a newly identified target transcript, which encodes for a plastid-



localised PsbP domain-OEC23 like protein, was detected with decreased mRNA levels in the overexpression line compared to the wild type. The mRNA levels of the putative target were reduced by half, which suggest a cleavage mediated through miR156h. This target coding for a PsbP domain is necessary for stabilisation as well as activation of the cofactors of the oxygen-evolving complex (OEC) in the photosystem II (Gregor and Britt, 2000). As a result, miR156h is not only able to regulate plant development and flowering time by modulating transcripts of the *SPL* transcription factor family, but it might be also involved in PSII regulation by cleaving the mRNA transcript encoding for a plastid-localised PsbP domain-OEC23 like protein. As a positive control for miR398b, transcript levels of the validated target *CSD2* have been detected to be decreased compared to the wild type. We analysed two additionally predicted targets for miR398b and we discovered that the transcript encoding the MATE protein might not get cleaved through miR398b, since the transcript levels were not decreased in the overexpression line of miR398b. However, the second putative transcript coding for a *PGM* protein showed reduced expression levels. The enzyme PGM catalyses the conversion of glucose-6-phosphate into glucose-1-phosphate at the beginning of the starch biosynthesis, an important pathway to produce and store energy in the plant cells (Zhang et al., 2019). Interestingly, *Arabidopsis thaliana* mutants defective in the starch biogenesis pathway have negative impacts in the cell wall metabolism and structure (Engelsdorf et al., 2017). Due to reduced transcript levels of *PGM* in the overexpression line of miR398b, we can assume that miR398b is able to mediate the cleavage of the mRNA transcript encoding for *PGM*. Consequently, the overexpression line of miR398b should develop a phenotype in case miR398b is able to mediate the cleavage of the *PGM* transcript. The validated target, *IAA28*, was not detected with remarkable decreased expression levels in the miR847 overexpression line compared to the wild type. We checked three putative targets that show a considerable downregulation in expression values in the miR847 overexpression line. The first miRNA target encodes for a PPR protein, which was already identified as *LOW PHOTOSYNTHETIC EFFICIENCY 1 (LPE1)* and is able to bind the mRNA of *PHOTOSYSTEM II REACTION CENTER PROTEIN A (PSBA)*, encoding the D1

protein of the photosystem II reaction centre (Jin et al., 2018). The second putative target of miR847 encodes for the *GENERAL CONTROL NON-REPRESSIBLE 2* (*GCN2*) kinase, which binds uncharged tRNA and hence supports the translation initiation factor eIF2 $\alpha$  (Lokdarshi et al., 2020). The third interesting putative target identified with decreased expression levels in the miR847 overexpression line compared to the wild type, encodes for the *PALE CRESS* (*PAC*) protein. *PAC* is known to play an important role in leaf and chloroplast developmental stages and the albinotic *pac* mutant has very drastic translational deficiencies (Meurer et al., 2017). Briefly, these findings lead to the assumption that miR847 targets the mRNA transcripts, which can be important in the development and stress acclimation.

In summary, we demonstrated that sRNAs play a role in response to NF and additionally under low temperatures. Furthermore, changes in the steady-state expression levels of sRNAs can be triggered by environmental stimuli such as cold or plastid derived signals. In both publications most of the differentially regulated sRNAs belong to the class of *cis*- and *trans*-nat-siRNAs followed by miRNAs. Therefore, we hypothesise that both sRNA classes, nat-siRNAs and miRNAs, are important players, which regulate mRNA transcript in response to NF or low temperatures. Besides sRNAs, we detected a high number of differentially regulated nuclear encoded genes, which are not related to plastid derived signals until now. Subsequently, we combined mRNA/lncRNA sequencing together with sRNA sequencing and correlated differentially expressed sRNAs with putative RNA targets to detect sRNA:RNA target pairs. Those sRNA:RNA target pairs might operate in the adaptation of plastidic and nuclear gene expression in retrograde signalling pathways or in the regulatory networks in cold acclimation. We found that different miRNAs have been identified to play a role in different developmental stages or have been known to be important in response to stress conditions. Nevertheless, further investigations, like RACE-PCR and proteomic experiments on the protein amount of the miRNA targets on the partially analysed miRNA overexpression lines (miR156h, miR398b and miR847) and additional miRNA overexpression lines

(e.g. miR157a or miR854) are needed to uncover the importance of sRNAs in stress response. Besides overexpression, mimicry lines can help in studying functions of miRNAs, because they have modified versions of the lncRNA IPS, which acts as a sequester for miRNAs (Franco-Zorrilla et al., 2007). Available miRNA mimicry lines like MIM156, MIM157, MIM169efg, MIM395, MIM398, MIM847 (Todesco et al., 2010) can help to uncover new potential roles in retrograde signalling or cold responsive pathways.

## REFERENCES

- ADHIKARI, N. D., FROEHLICH, J. E., STRAND, D. D., BUCK, S. M., KRAMER, D. M. & LARKIN, R. M. 2011. GUN4-porphyrin complexes bind the ChlH/GUN5 subunit of Mg-Chelatase and promote chlorophyll biosynthesis in Arabidopsis. *Plant Cell*, 23, 1449-1467.
- ALLEN, E., XIE, Z., GUSTAFSON, A. M. & CARRINGTON, J. C. 2005. microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*, 121, 207-221.
- AUNG, K., LIN, S. I., WU, C. C., HUANG, Y. T., SU, C. L. & CHIOU, T. J. 2006. *pho2*, a phosphate overaccumulator, is caused by a nonsense mutation in a microRNA399 target gene. *Plant Physiol*, 141, 1000-1011.
- AXTELL, M. J. 2013. Classification and comparison of small RNAs from plants. *Annu Rev Plant Biol*, 64, 137-159.
- AXTELL, M. J., JAN, C., RAJAGOPALAN, R. & BARTEL, D. P. 2006. A two-hit trigger for siRNA biogenesis in plants. *Cell*, 127, 565-577.
- BAEK, D., CHUN, H. J., KANG, S., SHIN, G., PARK, S. J., HONG, H., KIM, C., KIM, D. H., LEE, S. Y., KIM, M. C. & YUN, D. J. 2016. A role for Arabidopsis miR399f in salt, drought, and ABA signaling. *Mol Cells*, 39, 111-118.
- BALL, L., ACCOTTO, G. P., BECHTOLD, U., CREISSEN, G., FUNCK, D., JIMENEZ, A., KULAR, B., LEYLAND, N., MEJIA-CARRANZA, J., REYNOLDS, H., KARPINSKI, S. & MULLINEAUX, P. M. 2004. Evidence for a direct link between glutathione biosynthesis and stress defense gene expression in Arabidopsis. *Plant Cell*, 16, 2448-2462.
- BANNISTER, A. J. & KOUZARIDES, T. 2011. Regulation of chromatin by histone modifications. *Cell Res*, 21, 381-395.
- BARDOU, F., ARIEL, F., SIMPSON, C. G., ROMERO-BARRIOS, N., LAPORTE, P., BALZERGUE, S., BROWN, J. W. & CRESPI, M. 2014. Long noncoding RNA modulates alternative splicing regulators in Arabidopsis. *Dev Cell*, 30, 166-176.
- BARI, A., ORAZOVA, S. & IVASHCHENKO, A. 2013. miR156- and miR171-binding sites in the protein-coding sequences of several plant genes. *Biomed Res Int*, 2013, 307145.
- BARKAN, A. & SMALL, I. 2014. Pentatricopeptide repeat proteins in plants. *Annu Rev Plant Biol*, 65, 415-442.
- BAULINA, N. M., KULAKOVA, O. G. & FAVOROVA, O. O. 2016. MicroRNAs: the role in autoimmune inflammation. *Acta Naturae*, 8, 21-33.
- BEN-GERA, H., DAFNA, A., ALVAREZ, J. P., BAR, M., MAUERER, M. & ORI, N. 2016. Auxin-mediated lamina growth in tomato leaves is restricted by two parallel mechanisms. *Plant J*, 86, 443-457.
- BERRY, J. O., YERRAMSETTY, P., ZIELINSKI, A. M. & MURE, C. M. 2013. Photosynthetic gene expression in higher plants. *Photosynth Res*, 117, 91-120.
- BOBIK, K. & BURCH-SMITH, T. M. 2015. Chloroplast signaling within, between and beyond cells. *Front Plant Sci*, 6, 781.

- BORGES, F. & MARTIENSSEN, R. A. 2015. The expanding world of small RNAs in plants. *Nat Rev Mol Cell Biol*, 16, 727-741.
- BORSANI, O., ZHU, J., VERSLUES, P. E., SUNKAR, R. & ZHU, J. K. 2005. Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. *Cell*, 123, 1279-1291.
- CATALANOTTO, C., COGONI, C. & ZARDO, G. 2016. MicroRNA in control of gene expression: an overview of nuclear functions. *Int J Mol Sci*, 17, 1712.
- CHINNUSAMY, V., ZHU, J. & ZHU, J. K. 2007. Cold stress regulation of gene expression in plants. *Trends Plant Sci*, 12, 444-4451.
- CHINNUSAMY, V., ZHU, J. K. & SUNKAR, R. 2010. Gene regulation during cold stress acclimation in plants. *Methods Mol Biol*, 639, 39-55.
- CHIOU, T. J., AUNG, K., LIN, S. I., WU, C. C., CHIANG, S. F. & SU, C. L. 2006. Regulation of phosphate homeostasis by microRNA in Arabidopsis. *Plant Cell*, 18, 412-421.
- CHOI, M. & DAVIDSON, V. L. 2011. Cupredoxins--a study of how proteins may evolve to use metals for bioenergetic processes. *Metallomics*, 3, 140-151.
- CRISP, P. A., GANGULY, D. R., SMITH, A. B., MURRAY, K. D., ESTAVILLO, G. M., SEARLE, I., FORD, E., BOGDANOVIC, O., LISTER, R., BOREVITZ, J. O., EICHEN, S. R. & POGSON, B. J. 2017. Rapid recovery gene downregulation during excess-light stress and recovery in Arabidopsis. *Plant Cell*, 29, 1836-1863.
- CUI, L. G., SHAN, J. X., SHI, M., GAO, J. P. & LIN, H. X. 2014. The miR156-SPL9-DFR pathway coordinates the relationship between development and abiotic stress tolerance in plants. *Plant J*, 80, 1108-1117.
- D'ALESSANDRO, S., KSAS, B. & HAVAUX, M. 2018. Decoding beta-cyclocitral-mediated retrograde signaling reveals the role of a detoxification response in plant tolerance to photooxidative stress. *Plant Cell*, 30, 2495-2511.
- DEVERS, E. A., BRANSCHIED, A., MAY, P. & KRAJINSKI, F. 2011. Stars and symbiosis: microRNA- and microRNA\*-mediated transcript cleavage involved in arbuscular mycorrhizal symbiosis. *Plant Physiol*, 156, 1990-2010.
- DIETZ, K. J., TURKAN, I. & KRIEGER-LISZKAY, A. 2016. Redox- and reactive oxygen species-dependent signaling into and out of the photosynthesizing chloroplast. *Plant Physiol*, 171, 1541-1550.
- DING, Y., SHI, Y. & YANG, S. 2019. Advances and challenges in uncovering cold tolerance regulatory mechanisms in plants. *New Phytol*, 222, 1690-1704.
- DONG, C., PEI, H. 2014. Over-expression of miR397 improves plant tolerance to cold stress in Arabidopsis thaliana. *J Plant Biol*, 57, 209-217.
- ELLIS, J., DODDS, P. & PRYOR, T. 2000. Structure, function and evolution of plant disease resistance genes. *Curr Opin Plant Biol*, 3, 278-284.
- ENGELSDORF, T., WILL, C., HOFMANN, J., SCHMITT, C., MERRITT, B. B., RIEGER, L., FRENGER, M. S., MARSCHALL, A., FRANKE, R. B., PATTATHIL, S. & VOLL, L. M. 2017. Cell wall composition and penetration resistance against the fungal pathogen Colletotrichum higginsianum are affected by impaired starch turnover in Arabidopsis mutants. *J Exp Bot*, 68, 701-713.
- ESTAVILLO, G. M., CRISP, P. A., PORNISIRIWONG, W., WIRTZ, M., COLLINGE, D., CARRIE, C., GIRAUD, E., WHELAN, J., DAVID, P.,

- JAVOT, H., BREARLEY, C., HELL, R., MARIN, E. & POGSON, B. J. 2011. Evidence for a SAL1-PAP chloroplast retrograde pathway that functions in drought and high light signaling in Arabidopsis. *Plant Cell*, 23, 3992-4012.
- FAHLGREN, N., MONTGOMERY, T. A., HOWELL, M. D., ALLEN, E., DVORAK, S. K., ALEXANDER, A. L. & CARRINGTON, J. C. 2006. Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA affects developmental timing and patterning in Arabidopsis. *Curr Biol*, 16, 939-944.
- FANG, X., ZHAO, G., ZHANG, S., LI, Y., GU, H., LI, Y., ZHAO, Q. & QI, Y. 2018. Chloroplast-to-nucleus signaling regulates microRNA biogenesis in Arabidopsis. *Dev Cell*, 48, 371-382.
- FENG, P., GUO, H., CHI, W., CHAI, X., SUN, X., XU, X., MA, J., ROCHAIX, J. D., LEISTER, D., WANG, H., LU, C. & ZHANG, L. 2016. Chloroplast retrograde signal regulates flowering. *Proc Natl Acad Sci U S A*, 113, 10708-10713.
- FERNANDO, V. C. D., AL KHATEEB, W., BELMONTE, M. F. & SCHROEDER, D. F. 2018. Role of Arabidopsis ABF1/3/4 during det1 germination in salt and osmotic stress conditions. *Plant Mol Biol*, 97, 149-163.
- FRANCO-ZORRILLA, J. M., VALLI, A., TODESCO, M., MATEOS, I., PUGA, M. I., RUBIO-SOMOZA, I., LEYVA, A., WEIGEL, D., GARCIA, J. A. & PAZ-ARES, J. 2007. Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat Genet*, 39, 1033-1037.
- GAO, R., WANG, Y., GRUBER, M. Y. & HANNOUFA, A. 2017. miR156/SPL10 modulates lateral root development, branching and leaf morphology in Arabidopsis by silencing AGAMOUS-LIKE 79. *Front Plant Sci*, 8, 2226.
- GARCIA-MOLINA, A., KLEINE, T., SCHNEIDER, K., MUHLHAUS, T., LEHMANN, M. & LEISTER, D. 2020. Translational components contribute to acclimation responses to high light, heat, and cold in Arabidopsis. *iScience*, 23, 101331.
- GOLLAN, P. J., TIKKANEN, M. & ARO, E. M. 2015. Photosynthetic light reactions: integral to chloroplast retrograde signalling. *Curr Opin Plant Biol*, 27, 180-191.
- GRAY, J. C., SULLIVAN, J. A., WANG, J. H., JEROME, C. A. & MACLEAN, D. 2003. Coordination of plastid and nuclear gene expression. *Philos Trans R Soc Lond B Biol Sci*, 358, 135-145.
- GRAY, M. W., BURGER, G. & LANG, B. F. 1999. Mitochondrial evolution. *Science*, 283, 1476-1481.
- GREGOR, W. & BRITT, R. D. 2000. Nitrogen ligation to the manganese cluster of photosystem II in the absence of the extrinsic proteins and as a function of pH. *Photosynth Res*, 65, 175-185.
- GUTIERREZ-NAVA, M. D. L., GILLMOR, C. S., JIMENEZ, L. F., GUEVARA-GARCIA, A. & LEON, P. 2004. CHLOROPLAST BIOGENESIS genes act cell and noncell autonomously in early chloroplast development. *Plant Physiol*, 135, 471-482.
- HABERMANN, K., TIWARI, B., KRANTZ, M., ADLER, S. O., KLIPP, E., ARIF, M. A. & FRANK, W. 2020. Identification of small non-coding RNAs responsive to GUN1 and GUN5 related retrograde signals in Arabidopsis thaliana. *Plant J*, 104, 138-155.

- HAJYZADEH, M., TURKTAS, M., KHAWAR, K. M. & UNVER, T. 2015. miR408 overexpression causes increased drought tolerance in chickpea. *Gene*, 555, 186-193.
- HE, J., XU, M., WILLMANN, M. R., MCCORMICK, K., HU, T., YANG, L., STARKER, C. G., VOYTAS, D. F., MEYERS, B. C. & POETHIG, R. S. 2018. Threshold-dependent repression of SPL gene expression by miR156/miR157 controls vegetative phase change in *Arabidopsis thaliana*. *PLoS Genet*, 14, e1007337.
- HEO, J. B. & SUNG, S. 2011. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science*, 331, 76-79.
- HOLOCH, D. & MOAZED, D. 2015. RNA-mediated epigenetic regulation of gene expression. *Nat Rev Genet*, 16, 71-84.
- HUANG, C. Y., WANG, H., HU, P., HAMBY, R. & JIN, H. 2019. Small RNAs - big players in plant-microbe interactions. *Cell Host Microbe*, 26, 173-182.
- HUANG, G. T., MA, S. L., BAI, L. P., ZHANG, L., MA, H., JIA, P., LIU, J., ZHONG, M. & GUO, Z. F. 2012. Signal transduction during cold, salt, and drought stresses in plants. *Mol Biol Rep*, 39, 969-987.
- JIN, H., FU, M., DUAN, Z., DUAN, S., LI, M., DONG, X., LIU, B., FENG, D., WANG, J., PENG, L. & WANG, H. B. 2018. LOW PHOTOSYNTHETIC EFFICIENCY 1 is required for light-regulated photosystem II biogenesis in *Arabidopsis*. *Proc Natl Acad Sci U S A*, 115, E6075-E6084.
- KACPRZAK, S. M., MOCHIZUKI, N., NARANJO, B., XU, D., LEISTER, D., KLEINE, T., OKAMOTO, H. & TERRY, M. J. 2019. Plastid-to-nucleus retrograde signalling during chloroplast biogenesis does not require ABI4. *Plant Physiol*, 179, 18-23.
- KATIYAR-AGARWAL, S., MORGAN, R., DAHLBECK, D., BORSANI, O., VILLEGAS, A., JR., ZHU, J. K., STASKAWICZ, B. J. & JIN, H. 2006. A pathogen-inducible endogenous siRNA in plant immunity. *Proc Natl Acad Sci U S A*, 103, 18002-18007.
- KAZEMI-SHAHANDASHTI, S. S. & MAALI-AMIRI, R. 2018. Global insights of protein responses to cold stress in plants: signaling, defence, and degradation. *J Plant Physiol*, 226, 123-135.
- KHRAIWESH, B., ARIF, M. A., SEUMEL, G. I., OSSOWSKI, S., WEIGEL, D., RESKI, R. & FRANK, W. 2010. Transcriptional control of gene expression by microRNAs. *Cell*, 140, 111-122.
- KHRAIWESH, B., ZHU, J. K. & ZHU, J. 2012. Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *Biochim Biophys Acta*, 1819, 137-148.
- KIM, C. 2019. ROS-driven oxidative modification: its impact on chloroplasts-nucleus communication. *Front Plant Sci*, 10, 1729.
- KIM, C. & APEL, K. 2013. 1O<sub>2</sub>-mediated and EXECUTER-dependent retrograde plastid-to-nucleus signaling in norflurazon-treated seedlings of *Arabidopsis thaliana*. *Mol Plant*, 6, 1580-1591.
- KIM, J. Y., LEE, H. J., JUNG, H. J., MARUYAMA, K., SUZUKI, N. & KANG, H. 2010. Overexpression of microRNA395c or 395e affects differently the seed germination of *Arabidopsis thaliana* under stress conditions. *Planta*, 232, 1447-1454.



- KLEIN, M. & PAPENBROCK, J. 2004. The multi-protein family of Arabidopsis sulphotransferases and their relatives in other plant species. *J Exp Bot*, 55, 1809-1820.
- KLEINE, T. & LEISTER, D. 2016. Retrograde signaling: organelles go networking. *Biochim Biophys Acta*, 1857, 1313-1325.
- KLEINE, T., VOIGT, C. & LEISTER, D. 2009. Plastid signalling to the nucleus: messengers still lost in the mists? *Trends Genet*, 25, 185-192.
- KOORNNEEF, M. & VAN DER VEEN, J. H. 1980. Induction and analysis of gibberellin sensitive mutants in Arabidopsis thaliana (L.) heynh. *Theor Appl Genet*, 58, 257-263.
- KOUSSEVITZKY, S., NOTT, A., MOCKLER, T. C., HONG, F., SACHETTO-MARTINS, G., SURPIN, M., LIM, J., MITTLER, R. & CHORY, J. 2007. Signals from chloroplasts converge to regulate nuclear gene expression. *Science*, 316, 715-719.
- KRIMMER, T., RASSOW, J., KUNAU, W. H., VOOS, W. & PFANNER, N. 2000. Mitochondrial protein import motor: the ATPase domain of matrix Hsp70 is crucial for binding to Tim44, while the peptide binding domain and the carboxy-terminal segment play a stimulatory role. *Mol Cell Biol*, 20, 5879-5887.
- LAPIDOT, M. & PILPEL, Y. 2006. Genome-wide natural antisense transcription: coupling its regulation to its different regulatory mechanisms. *EMBO Rep*, 7, 1216-1222.
- LARKIN, R. M., ALONSO, J. M., ECKER, J. R. & CHORY, J. 2003. GUN4, a regulator of chlorophyll synthesis and intracellular signaling. *Science*, 299, 902-906.
- LAUBINGER, S., SACHSENBERG, T., ZELLER, G., BUSCH, W., LOHMANN, J. U., RATSCH, G. & WEIGEL, D. 2008. Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in Arabidopsis thaliana. *Proc Natl Acad Sci U S A*, 105, 8795-8800.
- LEE, S. J., KANG, J. Y., PARK, H. J., KIM, M. D., BAE, M. S., CHOI, H. I. & KIM, S. Y. 2010. DREB2C interacts with ABF2, a bZIP protein regulating abscisic acid-responsive gene expression, and its overexpression affects abscisic acid sensitivity. *Plant Physiol*, 153, 716-727.
- LEISTER, D. & KLEINE, T. 2016. Definition of a core module for the nuclear retrograde response to altered organellar gene expression identifies GLK overexpressors as gun mutants. *Physiol Plant*, 157, 297-309.
- LEISTER, D., WANG, L. & KLEINE, T. 2017. Organellar gene expression and acclimation of plants to environmental stress. *Front Plant Sci*, 8, 387.
- LI, F., PIGNATTA, D., BENDIX, C., BRUNKARD, J. O., COHN, M. M., TUNG, J., SUN, H., KUMAR, P. & BAKER, B. 2012. MicroRNA regulation of plant innate immune receptors. *Proc Natl Acad Sci U S A*, 109, 1790-1795.
- LI, S., CASTILLO-GONZALEZ, C., YU, B. & ZHANG, X. 2017. The functions of plant small RNAs in development and in stress responses. *Plant J*, 90, 654-670.
- LI, W. X., OONO, Y., ZHU, J., HE, X. J., WU, J. M., IIDA, K., LU, X. Y., CUI, X., JIN, H. & ZHU, J. K. 2008. The Arabidopsis NFYA5 transcription factor is

- regulated transcriptionally and posttranscriptionally to promote drought resistance. *Plant Cell*, 20, 2238-2251.
- LIANG, G., YANG, F. & YU, D. 2010. MicroRNA395 mediates regulation of sulfate accumulation and allocation in *Arabidopsis thaliana*. *Plant J*, 62, 1046-1057.
- LISTER, R., CHEW, O., LEE, M. N., HEAZLEWOOD, J. L., CLIFTON, R., PARKER, K. L., MILLAR, A. H. & WHELAN, J. 2004. A transcriptomic and proteomic characterization of the *Arabidopsis* mitochondrial protein import apparatus and its response to mitochondrial dysfunction. *Plant Physiol*, 134, 777-789.
- LIU, H. H., TIAN, X., LI, Y. J., WU, C. A. & ZHENG, C. C. 2008. Microarray-based analysis of stress-regulated microRNAs in *Arabidopsis thaliana*. *RNA*, 14, 836-843.
- LIU, P. P., MONTGOMERY, T. A., FAHLGREN, N., KASSCHAU, K. D., NONOGAKI, H. & CARRINGTON, J. C. 2007. Repression of AUXIN RESPONSE FACTOR10 by microRNA160 is critical for seed germination and post-germination stages. *Plant J*, 52, 133-146.
- LIU, X., HAO, L., LI, D., ZHU, L. & HU, S. 2015. Long non-coding RNAs and their biological roles in plants. *Genomics Proteomics Bioinformatics*, 13, 137-147.
- LLAMAS, E., PULIDO, P. & RODRIGUEZ-CONCEPCION, M. 2017. Interference with plastome gene expression and Clp protease activity in *Arabidopsis* triggers a chloroplast unfolded protein response to restore protein homeostasis. *PLoS Genet*, 13, e1007022.
- LOKDARSHI, A., GUAN, J., URQUIDI CAMACHO, R. A., CHO, S. K., MORGAN, P. W., LEONARD, M., SHIMONO, M., DAY, B. & VON ARNIM, A. G. 2020. Light activates the translational regulatory kinase GCN2 via reactive oxygen species emanating from the chloroplast. *Plant Cell*, 32, 1161-1178.
- LUO, Q. J., MITTAL, A., JIA, F. & ROCK, C. D. 2012. An autoregulatory feedback loop involving PAP1 and TAS4 in response to sugars in *Arabidopsis*. *Plant Mol Biol*, 80, 117-129.
- MA, C., BURD, S. & LERS, A. 2015. miR408 is involved in abiotic stress responses in *Arabidopsis*. *Plant J*, 84, 169-187.
- MA, X., SHAO, C., JIN, Y., WANG, H. & MENG, Y. 2014. Long non-coding RNAs: a novel endogenous source for the generation of dicer-like 1-dependent small RNAs in *Arabidopsis thaliana*. *RNA Biol*, 11, 373-390.
- MANAVELLA, P. A., YANG, S. W. & PALATNIK, J. 2019. Keep calm and carry on: miRNA biogenesis under stress. *Plant J*, 99, 832-843.
- MARGULIS, L. 1970. Origin of eukaryotic cells: evidence and research implications for a theory of the origin and evolution of microbial, plant and animal cells on the precambrian earth. *Yale University Press*, 301-329.
- MARINO, G., NARANJO, B., WANG, J., PENZLER, J. F., KLEINE, T. & LEISTER, D. 2019. Relationship of GUN1 to FUG1 in chloroplast protein homeostasis. *Plant J*, 99, 521-535.
- MARTIN, G., LEIVAR, P., LUDEVID, D., TEPPERMAN, J. M., QUAIL, P. H. & MONTE, E. 2016. Phytochrome and retrograde signalling pathways converge to antagonistically regulate a light-induced transcriptional network. *Nat Commun*, 7, 11431.

- MARTIN, W. & KOWALLIK, K. 1999. Annotated english translation of Mereschkowsky's 1905 paper 'Über Natur und Ursprung der Chromatophoren im Pflanzenreiche'. *European Journal of Phycology*, 34, 287-295.
- MARTIN, W., RUJAN, T., RICHLY, E., HANSEN, A., CORNELSEN, S., LINS, T., LEISTER, D., STOEBE, B., HASEGAWA, M. & PENNY, D. 2002. Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc Natl Acad Sci U S A*, 99, 12246-12251.
- MEGHA, S., BASU, U. & KAV, N. N. V. 2018. Regulation of low temperature stress in plants by microRNAs. *Plant Cell Environ*, 41, 1-15.
- MENG, Y., SHAO, C., MA, X., WANG, H. & CHEN, M. 2012. Expression-based functional investigation of the organ-specific microRNAs in Arabidopsis. *PLoS One*, 7, e50870.
- MESKAUSKIENE, R., NATER, M., GOSLINGS, D., KESSLER, F., OP DEN CAMP, R. & APEL, K. 2001. FLU: a negative regulator of chlorophyll biosynthesis in Arabidopsis thaliana. *Proc Natl Acad Sci U S A*, 98, 12826-12831.
- MEURER, J., SCHMID, L. M., STOPPEL, R., LEISTER, D., BRACHMANN, A. & MANAVSKI, N. 2017. PALE CRESS binds to plastid RNAs and facilitates the biogenesis of the 50S ribosomal subunit. *Plant J*, 92, 400-413.
- MICHAELS, S. D. & AMASINO, R. M. 1999. FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell*, 11, 949-956.
- MOCHIZUKI, N., BRUSSLAN, J. A., LARKIN, R., NAGATANI, A. & CHORY, J. 2001. Arabidopsis genomes uncoupled 5 (GUN5) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. *Proc Natl Acad Sci USA*, 98, 2053-2058.
- MOCHIZUKI, N., TANAKA, R., TANAKA, A., MASUDA, T. & NAGATANI, A. 2008. The steady-state level of Mg-protoporphyrin IX is not a determinant of plastid-to-nucleus signaling in Arabidopsis. *Proc Natl Acad Sci U S A*, 105, 15184-15189.
- MURCHA, M. W., LISTER, R., HO, A. Y. & WHELAN, J. 2003. Identification, expression, and import of components 17 and 23 of the inner mitochondrial membrane translocase from Arabidopsis. *Plant Physiol*, 131, 1737-1747.
- NAVARRO, L., DUNOYER, P., JAY, F., ARNOLD, B., DHARMASIRI, N., ESTELLE, M., VOINNET, O. & JONES, J. D. 2006. A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. *Science*, 312, 436-439.
- NIU, D., LII, Y. E., CHELLAPPAN, P., LEI, L., PERALTA, K., JIANG, C., GUO, J., COAKER, G. & JIN, H. 2016. miRNA863-3p sequentially targets negative immune regulator ARLPKs and positive regulator SERRATE upon bacterial infection. *Nat Commun*, 7, 11324.
- OELMULLER, R. & MOHR, H. 1986. Photooxidative destruction of chloroplasts and its consequences for expression of nuclear genes. *Planta*, 167, 106-113.

- PAGE, M. T., KACPRZAK, S. M., MOCHIZUKI, N., OKAMOTO, H., SMITH, A. G. & TERRY, M. J. 2017. Seedlings lacking the PTM protein do not show a genomes uncoupled (gun) mutant phenotype. *Plant Physiol*, 174, 21-26.
- PARK, W., LI, J., SONG, R., MESSING, J. & CHEN, X. 2002. CARPEL FACTORY, a dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. *Curr Biol*, 12, 1484-1495.
- PEGLER, J. L., OULTRAM, J. M. J., GROF, C. P. L. & EAMENS, A. L. 2019. Profiling the abiotic stress responsive microRNA landscape of *Arabidopsis thaliana*. *Plants (Basel)*, 8, 58.
- PESARESI, P. & KIM, C. 2019. Current understanding of GUN1: a key mediator involved in biogenic retrograde signaling. *Plant Cell Rep*, 38, 819-823.
- RAJAGOPALAN, R., VAUCHERET, H., TREJO, J. & BARTEL, D. P. 2006. A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. *Genes Dev*, 20, 3407-3425.
- RAJWANSHI, R., CHAKRABORTY, S., JAYANANDI, K., DEB, B. & LIGHTFOOT, D. A. 2014. Orthologous plant microRNAs: microregulators with great potential for improving stress tolerance in plants. *Theor Appl Genet*, 127, 2525-2543.
- RAMEL, F., BIRTIC, S., GINIES, C., SOUBIGOU-TACONNAT, L., TRIANTAPHYLIDES, C. & HAVAUX, M. 2012. Carotenoid oxidation products are stress signals that mediate gene responses to singlet oxygen in plants. *Proc Natl Acad Sci U S A*, 109, 5535-5540.
- REINHART, B. J., WEINSTEIN, E. G., RHOADES, M. W., BARTEL, B. & BARTEL, D. P. 2002. MicroRNAs in plants. *Genes Dev*, 16, 1616-1626.
- REN, L. & TANG, G. 2012. Identification of sucrose-responsive microRNAs reveals sucrose-regulated copper accumulations in an SPL7-dependent and independent manner in *Arabidopsis thaliana*. *Plant Sci*, 187, 59-68.
- RICHTER, A. S., TOHGE, T., FERNIE, A. R. & GRIMM, B. 2020. The genomes uncoupled-dependent signalling pathway coordinates plastid biogenesis with the synthesis of anthocyanins. *Philos Trans R Soc Lond B Biol Sci*, 375, 20190403.
- ROBERT-SEILANIANTZ, A., MACLEAN, D., JIKUMARU, Y., HILL, L., YAMAGUCHI, S., KAMIYA, Y. & JONES, J. D. 2011. The microRNA miR393 re-directs secondary metabolite biosynthesis away from camalexin and towards glucosinolates. *Plant J*, 67, 218-231.
- ROGERS, K. & CHEN, X. 2013. Biogenesis, turnover, and mode of action of plant microRNAs. *Plant Cell*, 25, 2383-2399.
- ROSSEL, J. B., WALTER, P. B., HENDRICKSON, L., CHOW, W. S., POOLE, A., MULLINEAUX, P. M. & POGSON, B. J. 2006. A mutation affecting ASCORBATE PEROXIDASE 2 gene expression reveals a link between responses to high light and drought tolerance. *Plant Cell Environ*, 29, 269-281.
- SAINI, G., MESKAUSKIENE, R., PIJACKA, W., ROSZAK, P., SJOGREN, L. L., CLARKE, A. K., STRAUS, M. & APEL, K. 2011. 'happy on norflurazon' (hon) mutations implicate perturbation of plastid homeostasis with activating stress acclimatization and changing nuclear gene expression in norflurazon-treated seedlings. *Plant J*, 65, 690-702.

- SAITO, T. & SAETROM, P. 2012. Target gene expression levels and competition between transfected and endogenous microRNAs are strong confounding factors in microRNA high-throughput experiments. *Silence*, 3, 3.
- SASAKI, K., KIM, M. H., KANNO, Y., SEO, M., KAMIYA, Y. & IMAI, R. 2015. Arabidopsis cold shock domain protein 2 influences ABA accumulation in seed and negatively regulates germination. *Biochem Biophys Res Commun*, 456, 380-384.
- SHIMIZU, T., KACPRZAK, S. M., MOCHIZUKI, N., NAGATANI, A., WATANABE, S., SHIMADA, T., TANAKA, K., HAYASHI, Y., ARAI, M., LEISTER, D., OKAMOTO, H., TERRY, M. J. & MASUDA, T. 2019. The retrograde signaling protein GUN1 regulates tetrapyrrole biosynthesis. *Proc Natl Acad Sci U S A*, 116, 24900-24906.
- SHRIRAM, V., KUMAR, V., DEVARUMATH, R. M., KHARE, T. S. & WANI, S. H. 2016. MicroRNAs as potential targets for abiotic stress tolerance in plants. *Front Plant Sci*, 7, 817.
- SHUKLA, P. S., BORZA, T., CRITCHLEY, A. T., HILTZ, D., NORRIE, J. & PRITHIVIRAJ, B. 2018. Ascophyllum nodosum extract mitigates salinity stress in Arabidopsis thaliana by modulating the expression of miRNA involved in stress tolerance and nutrient acquisition. *PLoS One*, 13, e0206221.
- SIMPSON, C. L. & STERN, D. B. 2002. The treasure trove of algal chloroplast genomes. Surprises in architecture and gene content, and their functional implications. *Plant Physiol*, 129, 957-966.
- SINGH, A., GAUTAM, V., SINGH, S., SARKAR DAS, S., VERMA, S., MISHRA, V., MUKHERJEE, S. & SARKAR, A. K. 2018. Plant small RNAs: advancement in the understanding of biogenesis and role in plant development. *Planta*, 248, 545-558.
- SONG, J. B., GAO, S., SUN, D., LI, H., SHU, X. X. & YANG, Z. M. 2013. miR394 and LCR are involved in Arabidopsis salt and drought stress responses in an abscisic acid-dependent manner. *BMC Plant Biol*, 13, 210.
- SONG, J. B., HUANG, S. Q., DALMAY, T. & YANG, Z. M. 2012. Regulation of leaf morphology by microRNA394 and its target LEAF CURLING RESPONSIVENESS. *Plant Cell Physiol*, 53, 1283-1294.
- SONG, Z., ZHANG, L., WANG, Y., LI, H., LI, S., ZHAO, H. & ZHANG, H. 2017. Constitutive expression of miR408 improves biomass and seed yield in Arabidopsis. *Front Plant Sci*, 8, 2114.
- STIEF, A., ALTMANN, S., HOFFMANN, K., PANT, B. D., SCHEIBLE, W. R. & BAURLE, I. 2014. Arabidopsis miR156 regulates tolerance to recurring environmental stress through SPL transcription factors. *Plant Cell*, 26, 1792-1807.
- STRAND, A., ASAMI, T., ALONSO, J., ECKER, J. R. & CHORY, J. 2003. Chloroplast to nucleus communication triggered by accumulation of Mg-protoporphyrinIX. *Nature*, 421, 79-83.
- SUAREZ, J. V., BANKS, S., THOMAS, P. G. & DAY, A. 2014. A new F131V mutation in Chlamydomonas phytoene desaturase locates a cluster of norflurazon resistance mutations near the FAD-binding site in 3D protein models. *PLoS One*, 9, e99894.

- SULLIVAN, J. A. & GRAY, J. C. 1999. Plastid translation is required for the expression of nuclear photosynthesis genes in the dark and in roots of the pea *lip1* mutant. *Plant Cell*, 11, 901-910.
- SUN, X., FENG, P., XU, X., GUO, H., MA, J., CHI, W., LIN, R., LU, C. & ZHANG, L. 2011. A chloroplast envelope-bound PHD transcription factor mediates chloroplast signals to the nucleus. *Nat Commun*, 2, 477.
- SUNKAR, R., KAPOOR, A. & ZHU, J. K. 2006. Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in Arabidopsis is mediated by downregulation of miR398 and important for oxidative stress tolerance. *Plant Cell*, 18, 2051-2065.
- SUNKAR, R. & ZHU, J. K. 2004. Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. *Plant Cell*, 16, 2001-2019.
- SURPIN, M. & CHORY, J. 1997. The co-ordination of nuclear and organellar genome expression in eukaryotic cells. *Essays Biochem*, 32, 113-125.
- SUSEK, R. E., AUSUBEL, F. M. & CHORY, J. 1993. Signal transduction mutants of Arabidopsis uncouple nuclear CAB and RBCS gene expression from chloroplast development. *Cell*, 74, 787-799.
- SWIEZEWSKI, S., LIU, F., MAGUSIN, A. & DEAN, C. 2009. Cold-induced silencing by long antisense transcripts of an Arabidopsis polycomb target. *Nature*, 462, 799-802.
- TADINI, L., PESARESI, P., KLEINE, T., ROSSI, F., GULJAMOW, A., SOMMER, F., MUHLHAUS, T., SCHRODA, M., MASIERO, S., PRIBIL, M., ROTHBART, M., HEDTKE, B., GRIMM, B. & LEISTER, D. 2016. GUN1 controls accumulation of the plastid ribosomal protein S1 at the protein level and interacts with proteins involved in plastid protein homeostasis. *Plant Physiol*, 170, 1817-1830.
- TANG, Z., XU, M., CAI, J., MA, X., QIN, J. & MENG, Y. 2019. Transcriptome-wide identification and functional investigation of the RDR2- and DCL3-dependent small RNAs encoded by long non-coding RNAs in Arabidopsis thaliana. *Plant Signal Behav*, 14, 1616518.
- TAYLOR, N. L., RUDHE, C., HULETT, J. M., LITHGOW, T., GLASER, E., DAY, D. A., MILLAR, A. H. & WHELAN, J. 2003. Environmental stresses inhibit and stimulate different protein import pathways in plant mitochondria. *FEBS Lett*, 547, 125-130.
- TERRY, M. J. & SMITH, A. G. 2013. A model for tetrapyrrole synthesis as the primary mechanism for plastid-to-nucleus signaling during chloroplast biogenesis. *Front Plant Sci*, 4, 14.
- THATCHER, S. R., BURD, S., WRIGHT, C., LERS, A. & GREEN, P. J. 2015. Differential expression of miRNAs and their target genes in senescing leaves and siliques: insights from deep sequencing of small RNAs and cleaved target RNAs. *Plant Cell Environ*, 38, 188-200.
- TIMMIS, J. N., AYLIFFE, M. A., HUANG, C. Y. & MARTIN, W. 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet*, 5, 123-135.
- TIWARI, B., HABERMANN, K., ARIF, M. A., WEIL, H. L., GARCIA-MOLINA, A., KLEINE, T., MUHLHAUS, T. & FRANK, W. 2020. Identification of small RNAs during cold acclimation in Arabidopsis thaliana. *BMC Plant Biol*, 20, 298.

- TODESCO, M., RUBIO-SOMOZA, I., PAZ-ARES, J. & WEIGEL, D. 2010. A collection of target mimics for comprehensive analysis of microRNA function in *Arabidopsis thaliana*. *PLoS Genet*, 6, e1001031.
- TRINDADE, I., CAPITAO, C., DALMAY, T., FEVEREIRO, M. P. & SANTOS, D. M. 2010. miR398 and miR408 are up-regulated in response to water deficit in *Medicago truncatula*. *Planta*, 231, 705-716.
- VIDAL, E. A., MOYANO, T. C., KROUK, G., KATARI, M. S., TANURDZIC, M., MCCOMBIE, W. R., CORUZZI, G. M. & GUTIERREZ, R. A. 2013. Integrated RNA-seq and sRNA-seq analysis identifies novel nitrate-responsive genes in *Arabidopsis thaliana* roots. *BMC Genomics*, 14, 701.
- VOINNET, O. 2009. Origin, biogenesis, and activity of plant microRNAs. *Cell*, 136, 669-687.
- WANG, H. V. & CHEKANOVA, J. A. 2017. Long noncoding RNAs in plants. *Adv Exp Med Biol*, 1008, 133-154.
- WANG, J. J. & GUO, H. S. 2015. Cleavage of INDOLE-3-ACETIC ACID INDUCIBLE28 mRNA by microRNA847 upregulates auxin signaling to modulate cell proliferation and lateral organ growth in *Arabidopsis*. *Plant Cell*, 27, 574-590.
- WANG, J. W. 2014. Regulation of flowering time by the miR156-mediated age pathway. *J Exp Bot*, 65, 4723-4730.
- WATERS, M. T., WANG, P., KORKARIC, M., CAPPER, R. G., SAUNDERS, N. J. & LANGDALE, J. A. 2009. GLK transcription factors coordinate expression of the photosynthetic apparatus in *Arabidopsis*. *Plant Cell*, 21, 1109-1128.
- WIERZBICKI, A. T., HAAG, J. R. & PIKAARD, C. S. 2008. Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell*, 135, 635-648.
- WOODSON, J. D., PEREZ-RUIZ, J. M. & CHORY, J. 2011. Heme synthesis by plastid ferrochelatase I regulates nuclear gene expression in plants. *Curr Biol*, 21, 897-903.
- WOODSON, J. D., PEREZ-RUIZ, J. M., SCHMITZ, R. J., ECKER, J. R. & CHORY, J. 2013. Sigma factor-mediated plastid retrograde signals control nuclear gene expression. *Plant J*, 73, 1-13.
- WU, G., PARK, M. Y., CONWAY, S. R., WANG, J. W., WEIGEL, D. & POETHIG, R. S. 2009. The sequential action of miR156 and miR172 regulates developmental timing in *Arabidopsis*. *Cell*, 138, 750-759.
- WU, G. & POETHIG, R. S. 2006. Temporal regulation of shoot development in *Arabidopsis thaliana* by miR156 and its target SPL3. *Development*, 133, 3539-3547.
- WU, G. Z., CHALVIN, C., HOELSCHER, M., MEYER, E. H., WU, X. N. & BOCK, R. 2018. Control of retrograde signaling by rapid turnover of GENOMES UNCOUPLED1. *Plant Physiol*, 176, 2472-2495.
- WU, M., SUN, L. V., VAMATHEVAN, J., RIEGLER, M., DEBOY, R., BROWNLIE, J. C., MCGRAW, E. A., MARTIN, W., ESSER, C., AHMADINEJAD, N., WIEGAND, C., MADUPU, R., BEANAN, M. J., BRINKAC, L. M., DAUGHERTY, S. C., DURKIN, A. S., KOLONAY, J. F., NELSON, W. C., MOHAMOUD, Y., LEE, P., BERRY, K., YOUNG, M. B., UTTERBACK, T., WEIDMAN, J., NIERMAN, W. C., PAULSEN, I. T., NELSON, K. E.,



- TETTELIN, H., O'NEILL, S. L. & EISEN, J. A. 2004. Phylogenomics of the reproductive parasite *Wolbachia pipientis* wMel: a streamlined genome overrun by mobile genetic elements. *PLoS Biol*, 2, E69.
- XU, M. Y., ZHANG, L., LI, W. W., HU, X. L., WANG, M. B., FAN, Y. L., ZHANG, C. Y. & WANG, L. 2014. Stress-induced early flowering is mediated by miR169 in *Arabidopsis thaliana*. *J Exp Bot*, 65, 89-101.
- YANG, B., SONG, Z., LI, C., JIANG, J., ZHOU, Y., WANG, R., WANG, Q., NI, C., LIANG, Q., CHEN, H. & FAN, L. M. 2018a. RSM1, an *Arabidopsis* MYB protein, interacts with HY5/HYH to modulate seed germination and seedling development in response to abscisic acid and salinity. *PLoS Genet*, 14, e1007839.
- YANG, T., WANG, Y., TEOTIA, S., ZHANG, Z. & TANG, G. 2018b. The making of leaves: how small RNA networks modulate leaf development. *Front Plant Sci*, 9, 824.
- YANG, Z., LI, W., SU, X., GE, P., ZHOU, Y., HAO, Y., SHU, H., GAO, C., CHENG, S., ZHU, G. & WANG, Z. 2019. Early response of radish to heat stress by strand-specific transcriptome and miRNA analysis. *Int J Mol Sci*, 20, 1-13.
- YOSHIKAWA, M. 2013. Biogenesis of trans-acting siRNAs, endogenous secondary siRNAs in plants. *Genes Genet Syst*, 88, 77-84.
- YU, B., YANG, Z., LI, J., MINAKHINA, S., YANG, M., PADGETT, R. W., STEWARD, R. & CHEN, X. 2005. Methylation as a crucial step in plant microRNA biogenesis. *Science*, 307, 932-935.
- YU, N., NIU, Q. W., NG, K. H. & CHUA, N. H. 2015. The role of miR156/SPLs modules in *Arabidopsis* lateral root development. *Plant J*, 83, 673-685.
- YU, Y., JIA, T. & CHEN, X. 2017. The 'how' and 'where' of plant microRNAs. *New Phytol*, 216, 1002-1017.
- YUAN, C., WANG, J., HARRISON, A. P., MENG, X., CHEN, D. & CHEN, M. 2015. Genome-wide view of natural antisense transcripts in *Arabidopsis thaliana*. *DNA Res*, 22, 233-243.
- ZHAI, J., JEONG, D. H., DE PAOLI, E., PARK, S., ROSEN, B. D., LI, Y., GONZALEZ, A. J., YAN, Z., KITTO, S. L., GRUSAK, M. A., JACKSON, S. A., STACEY, G., COOK, D. R., GREEN, P. J., SHERRIER, D. J. & MEYERS, B. C. 2011. MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. *Genes Dev*, 25, 2540-2553.
- ZHANG, J. P., YU, Y., FENG, Y. Z., ZHOU, Y. F., ZHANG, F., YANG, Y. W., LEI, M. Q., ZHANG, Y. C. & CHEN, Y. Q. 2017. MiR408 regulates grain yield and photosynthesis via a phytoeyanin protein. *Plant Physiol*, 175, 1175-1185.
- ZHANG, X., XIA, J., LII, Y. E., BARRERA-FIGUEROA, B. E., ZHOU, X., GAO, S., LU, L., NIU, D., CHEN, Z., LEUNG, C., WONG, T., ZHANG, H., GUO, J., LI, Y., LIU, R., LIANG, W., ZHU, J. K., ZHANG, W. & JIN, H. 2012. Genome-wide analysis of plant nat-siRNAs reveals insights into their distribution, biogenesis and function. *Genome Biol*, 13, R20.
- ZHANG, Y., HE, P., MA, X., YANG, Z., PANG, C., YU, J., WANG, G., FRIML, J. & XIAO, G. 2019. Auxin-mediated statolith production for root gravitropism. *New Phytol*, 224, 761-774.

- ZHANG, Z. W., WU, Z. L., FENG, L. Y., DONG, L. H., SONG, A. J., YUAN, M., CHEN, Y. E., ZENG, J., CHEN, G. D. & YUAN, S. 2016. Mg-Protoporphyrin IX Signals Enhance Plant's Tolerance to Cold Stress. *Front Plant Sci*, 7, 1545.
- ZHANG, Z. W., YUAN, S., XU, F., YANG, H., CHEN, Y. E., YUAN, M., XU, M. Y., XUE, L. W., XU, X. C. & LIN, H. H. 2011. Mg-protoporphyrin, haem and sugar signals double cellular total RNA against herbicide and high-light-derived oxidative stress. *Plant Cell Environ*, 34, 1031-1042.
- ZHAO, C., WANG, Y., CHAN, K. X., MARCHANT, D. B., FRANKS, P. J., RANDALL, D., TEE, E. E., CHEN, G., RAMESH, S., PHUA, S. Y., ZHANG, B., HILLS, A., DAI, F., XUE, D., GILLIHAM, M., TYERMAN, S., NEVO, E., WU, F., ZHANG, G., WONG, G. K., LEEBENS-MACK, J. H., MELKONIAN, M., BLATT, M. R., SOLTIS, P. S., SOLTIS, D. E., POGSON, B. J. & CHEN, Z. H. 2019a. Evolution of chloroplast retrograde signaling facilitates green plant adaptation to land. *Proc Natl Acad Sci U S A*, 116, 5015-5020.
- ZHAO, X., HUANG, J. & CHORY, J. 2019b. GUN1 interacts with MORF2 to regulate plastid RNA editing during retrograde signaling. *Proc Natl Acad Sci U S A*, 116, 10162-10167.
- ZHENG, C., YE, M., SANG, M. & WU, R. 2019. A Regulatory Network for miR156-SPL Module in *Arabidopsis thaliana*. *Int J Mol Sci*, 20.
- ZHU, J. K. 2016. Abiotic Stress Signaling and Responses in Plants. *Cell*, 167, 313-324.
- ZIMORSKI, V., KU, C., MARTIN, W. F. & GOULD, S. B. 2014. Endosymbiotic theory for organelle origins. *Curr Opin Microbiol*, 22, 38-48.

**ACKNOWLEDGMENTS**

First of all, I would like to thank Wolfgang Frank for giving me the opportunity to manage the Ph.D. project, but also for being a trusting and helpful supervisor and for the challenging tasks, endless support, understanding as well as the encouraging feedback during the whole process of this scientific work.

I would like to thank PD Dr. Tatjana Kleine as the “Zweitgutachter” for reviewing this work.

Also, many thanks to my TAC committee, Felix Willmund and Tatjana Kleine, for all the useful and conducive suggestions for the Ph.D. project.

Additionally, I would like to thank M. Asif Arif for giving me the freedom during experiments and the progress during the Ph.D. project as well as for helpful support whenever you had time.

Further, I would like to thank my PhD labmate Bhavika Tiwari for all the fruitful discussions and constant need of answering each other queries.

I am very grateful to the other colleagues from AG Frank like Stefan Kirchner and Oguz Top, who always shared their knowledge and gave psychological support. I thank my former research course students for providing me an opportunity to experience the teaching and learning process.

Additionally, I would like to thank also other colleagues from the botanical department. Special thanks to Dr. Giada Marino for inspiring conversations during lunch time and beyond that thanks for all the support during my PhD time as well as the elaboration of my thesis.

Thanks to the DFG for funding the project SFB-TRR175 and many thanks to all the collaboration partners and SFB members that helped me to interact with other members and gain more knowledge to other topics.

Last but not the least, I would like to thank my family and friends who supported me all these years.

**CURRICULUM VITAE**

<b>Name</b>	Kristin Habermann
<b>Date of birth</b>	09.01.1992
<b>Place of birth</b>	Lichtenstein, Germany

**Education**

Since 09.2016	PhD Student at Ludwig-Maximilians-Universität München Plant Molecular Cell Biology
09.2014 - 08.2016	Master of Science in Molekularbiologie/Bioinformatik at University of Applied Science in Mittweida Degree: Master of Science (1.6)
09.2011 - 09.2014	Bachelor of Science in Biotechnologie/Bioinformatik at University of Applied Science in Mittweida Degree: Bachelor of Science (2.3)
10.2010 - 03.2011	Study at TU Chemnitz in Wirtschaftsingenierwesen
09.2003 - 06.2010	Albert-Schweitzer-Gymnasium Limbach-Oberfrohna; Abitur (2.7)

**Awards**

09.2015 - 08.2016	Deutschlandstipendium
09.2014 - 08.2015	